现代数学丛书

潘一民 著

随机过程的线性统计理论与方法

LINEAR STATISTICAL THEORY AND METHODS IN STOCHASTIC PROCESSES

PAN YIMIN

·现代数学丛书·

随机过程的线性统计理论与方法

潘一民 著



上海科学技术出版社

责任编辑　　赵序明

现代数学丛书
随机过程的线性统计理论与方法
潘一民　著
上海科学技术出版社出版、发行
（上海瑞金二路450号）
新华书店上海发行所经销　　商务印书馆上海印刷厂印刷
开本 787×1092　1/20　印张 11.25　插页 4　字数 137,000
1993年11月第1版　　1993年11月第1次印刷
印数　1—1,400
ISBN 7-5323-3149-0/O·168
定价：18.00元

（沪）新登字108号

内 容 提 要

本书从再生核表示空间出发，统一而严格地论述了随机过程线性统计的理论基础，然后系统地介绍了各种线性随机系统的滤波、预测、平滑、参数估计以及最优控制的方法，特别是适用于计算机的递推算法，大大扩充了已知的模型。很多内容是作者本人的研究成果。本书可作为大学本科和研究生的参考书或教材，也可供有志于提高和深入的有关应用工作者参考。

Modern Mathematics Series

LINEAR STATISTICAL THEORY AND METHODS IN STOCHASTIC PROCESSES

Pan Yimin

Shanghai Scientific & Technical Publishers

Linear Statistical Theory and Methods in Stochastic Processes

Pan Yimin

Abstract

In this monograph, based on the reproducing kernel representation, we discussed the linear statistical theory of stochastic processes in an united and rigorous way, then systematically presented the methods, especially the recursive algorithms, of filtering, prediction, smoothing, estimation of parameters and optimal control for various linear stochastic systems, which extended the well-known linear systems with white-noise remarkably. Many results are due to the author.

《现代数学丛书》编辑委员会

Modern Mathematics Series
Editorial Committee

出版说明

从六十年代起，由华罗庚教授任主编的《现代数学丛书》编辑委员会曾组织编著，并由我社出版了多部具有很高水平的数学学术专著，有几部专著并已在国外出了外文版，受到国内外数学界和广大读者的高度重视，获得了很高的评价．原编委会中华罗庚、关肇直、吴新谋三位教授虽已先后逝世，但他们为本《丛书》所作出的贡献迄今仍为人们所敬仰、怀念．由于某些客观原因，《现代数学丛书》的出版工作曾一度停顿．

为了适应现代数学的迅速发展，更好地反映我国数学家近几年的优秀研究成果，必须大力加强《现代数学丛书》的规划、编辑、出版工作，充实编委会的力量．考虑到不少编委年事已高，经向原编委会中大部分同志及数学界有关专家广泛征求意见后，于 1990 年对编委会作了调整，补充了一些著名的中年数学家和学科带头人，建立了新的编委会，并进一步明确了本丛书的宗旨．

《现代数学丛书》新的编辑委员会由苏步青教授任名誉主编，谷超豪教授任主编，十八位著名数学家任委员．编委会负责推荐（或审定）选题和作者，主持书稿的审核等工作．

《现代数学丛书》的宗旨是：向国内外介绍我国比较成熟的、对学科发展方向有引导作用的、国内第一流水平的数学研究成果，反映我国数学研究的特色和优势，扩大我国数学研究成果的影响，促进学科的发展和国内外的学术交流．

为了实现上述宗旨，本丛书将陆续组织出版在基础数学、应用数学和计算数学方面处于学科发展前沿、有创见且具有系统完整

研究成果的现代数学学术专著.

为出版好《现代数学丛书》，我们热切地期望着数学界各位专家的大力支持和悉心指导，并欢迎广大读者提出宝贵的建议和意见。

上海科学技术出版社

前 言

一切统计问题的实质，都是要对随机现象(总体)的部分观测(样本)进行合理的分析，以求对该总体的性质作出某种判断、估计或预测. 如果所考虑的总体性质是随时间而演变的，就属于随机过程统计的范围. 在这类问题中，分析的观测值是前后相互关联的随机过程的一段或一部分采样(动态数据)，其结构比经典统计所分析的总体的简单重复抽样(静态数据)要复杂得多，因此研究手段和数学工具也相应地增加了难度. 其中方法上最成熟、应用上也最广泛的是随机过程的线性统计学，粗略地说，就是限于考虑由观测值的线性函数构成的统计量和均方最优准则. 适应这一条件的统计问题的范围固然有一定局限性，但也带来一个很大的好处，就是它们的最优解仅与过程的前二阶矩性质有关，而无须知道或假定总体的更详尽的概率分布性质. 从模型、理论与方法上讲，它是经典统计中相应问题的本质拓广，也是非线性统计手段的前提与基石；从应用角度讲，它概括了各种线性随机系统的滤波、预测、平滑、未知线性参数估计以及最优线性控制等内容. 据作者所知，迄今国内外还没有一本专门著作，既全面而严格地论述随机过程线性统计的理论基础，又统一而详尽地介绍解决各类有关问题的方法，特别是适用于计算机迭代运算的方法，而弥补这一缺陷，正是写作本书的一个目的.

此外，在过程统计的许多应用中，例如航天、雷达、电子对抗中的信号检测，人造卫星、火箭、工业过程的控制等等，动态数据并不是同时全部给出，而是按时间先后逐步到来的. 它们的数量极其

庞大，且需要适时地加以分析处理判断并作出必要的反应．这就对过程统计的方法论提出了严峻的要求。人们所期望的，往往不是涉及全部观测值的解析表达式，而是对动态数据的序贯处理．电子计算机的飞速进步和普遍使用更为这种要求提供了现实可能性．撰写本书的另一目的，就是希望对尽可能广泛的二阶随机过程(序列)和线性系统模型，提供一整套最优准则下的在线(on line)统计递推格式，使得在每增加一批观测数据时，只需依赖前面少数几步已有的估计值，便能进一步作出新的估计或改进，而无须保存历史数据，从而大大扩充著名的 Kalman 滤波方法．

本书是一本理论性著作，它包含了作者十多年来的研究工作(此项工作曾得到国家自然科学基金委员会资助)．许多结果是新的，而且大都是首次发表，某些已有的定理也给出了完全新的严格证明．第一章是全书的预备知识，即再生核表示理论．最先利用再生核空间来表述二阶过程及其线性统计问题的，乃是捷克数学家 J. Hájek[1]．在本章中，我们将这一概念推广到多维二阶过程，并明确地提出了正则过程这一重要定义．还给出了一套作者认为很方便的特殊记法．在第二章中我们详尽地介绍了任意指标集上随机过程的各类线性统计问题，论证了最优解存在的充分必要条件，并把它们统一归结为求出量测过程的再生核表示空间及其逆再生表示．其中有些问题是在参考文献[1]中讨论过的，这里只是把它们推广到多维情形．但还有一些问题，例如带线性约束条件的回归系数估计，关于可估性的讨论，以及注记 2.2.7 中含未知回归系数矩阵的模型，则是受经典统计模型的启发，第一次在本书中加以研究的．当然，作为特例，它们又包含了经典统计中对应问题的解答．

第三、四两章专门研究离散时间随机过程，即随机序列的线性统计问题．根据第二章的结果，我们首先讨论了在什么条件下，一个局部正则随机序列的再生核表示及逆再生表示可以进行有限步递推运算．为此，特别提出了“q 步后可分解协方差的随机序列”的概念．这一概念由作者与杜金观在参考文献[11]中首次给出，

本书又作了进一步推广. 我们证明了这类序列与一类广泛的非白噪声线性随机系统的输出系列是等价的,从而对后者的最优滤波、预测与平滑问题给出了一套完整的递推公式. 进一步，还研究了输入含有未知线性参数以及量测噪声为时变 ARMA 序列的系统,上述问题也获得圆满解决. 这些模型的范围已远远超出了通常处理的白噪声系统.

第五章论述随机积分与多维正交随机测度，它是为后两章研究连续时间随机过程作准备的,但是为了扩大它的内涵,我们在一般的阵测度空间上进行讨论. 本章的显著特点是用逆再生表示来定义两种随机积分，然后详细分析它们的基本性质. 书中还顺便指出了现代鞅论中关于局部有界可料过程对局部平方可积鞅的积分，可以容易地归结为本章定义的一种情形. 对平稳过程的谱表示问题也作了推广.

第六、七两章讨论连续时间过程的线性统计问题，其中的"线性新息定理"是一个十分重要的结果和关键的工具. 原苏联数学家 Y. A. Rozanov 曾为此专门写了一本小册子，证明的思路与方法都相当复杂. 我们用再生表示方法给出了一个简洁的新证明，并且减弱了原来的条件和推广到多维情形. 利用这一定理，我们严格推导出连续时间系统的 Kalman-Bucy 滤波公式，同时还与离散时间系统对应地研究了输入含未知参数的情形以及非积累量测的情形,并获得表为微分方程的滤波公式.

第八章的主题是线性随机系统的最优控制. 我们分别对离散时间与连续时间两种受控系统，给出了平均加权平方损失下最优控制的结构表达式,并与滤波公式作了对比. 结果虽然是经典的，但证明却有新意. 尤其是对连续时间系统，所给的证明既严格又初等,似乎尚未在别的文献中见到过.

本书的内容自成系统. 读者只需具备概率论与测度论的基础知识和对 Hilbert 空间理论的初步了解,就足以看懂全书(第五章 §5.4 可以略过). 对严格的理论推导和尽可能广泛的假设条件不感兴趣的读者,尤其是有志于提高和深入的有关应用工作者,也可

以在假定某些结果的前提下涉猎本书，找到对应用有启发的思想与方法．例如第四章列举的各类系统的递推滤波公式，第八章的最优控制结构，就可以很方便地编成计算机程序．

此外，本书也可以作为研究生的专业教材．

作者感谢老同学陈希孺教授，他于百忙之中仔细审阅了书稿，提出许多十分中肯的意见．感谢汪嘉冈教授，他一直关注我的研究工作，经常给予支持和鼓励．感谢上海科学技术出版社赵序明同志，他不惮其劳，热心推荐和编审，使本书得以迅速问世．最后，我还想把一片感激之情献给我的妻子，作为我的同事和同行，她自始至终倾注了极大的心力，没有她的帮助与促进，本书是难以写成的．

作者学识浅陋，如发现错误与不当之处，敬请同行与读者不吝指正．

潘一民

1992年9月于中国科学院应用数学研究所

记号说明

记号	意义	所在章节
T	任意非空指标集	1.1.1
N	正整数全体	1.1.1
C	复数(或实数)全体	1.1.1
$\mathscr{X}^k$	以集 $\mathscr{X}$ 的元为分量的 k 维列向量全体	1.1.1
$\mathscr{X}^{k\times l}$	以集 $\mathscr{X}$ 的元为元的 $k\times l$ 矩阵全体	1.1.1
$\triangleq$	表"定义为"或"记为"	1.1.1
C^*	复(或实)矩阵 C 的共轭转置(或转置)矩阵	1.1.1
$\Vert C\Vert$	复或实矩阵 C 的范数，$\Vert C\Vert\triangleq\sqrt{\mathrm{tr}(C^*C)}$	
	(其中 tr 表示方阵的迹)	1.1.1
Γ^-	复 Hermite 阵或实对称阵 Γ 的弱逆	1.1.1
$\Gamma\geqslant 0$	表 Γ 为半正定 Hermite 阵	1.1.1
$\Gamma_1\geqslant\Gamma_2$	表 $\Gamma_1-\Gamma_2\geqslant 0$	1.1.1
$(\Omega, \mathscr{F}, P)$	概率空间	1.1.1
$\mathscr{L}_2$	二阶矩有穷的随机变量全体	1.1.1
$\mathscr{L}_{20}$	二阶矩有穷且均值为 0 的随机变量全体	1.1.1
$R(x, y)$	随机向量 x 与 y 的协方差阵	1.1.1
$R(x)$	随机向量 x 的方差阵，$R(x)\triangleq R(x, x)$	1.1.1
$R^+(x, y)$	$\triangleq Exy^*$	1.1.1
$R^+(x)$	$\triangleq R^+(x, x)$	1.1.1
z_T	以 T 为指标集的(m 维)二阶随机过程	1.2.1
$\mathscr{H}_0(z_T)$	由 z_T 张成的线性空间	1.2.1
Z^x	由 z_T 与随机向量 x 确定的矩阵值函数，	

记号	说明	节
	$Z_t^x \triangleq R(x, z_t)$, $t\in T$	1.2.1
$\mathscr{H}(z_T)$	正则过程 z_T 生成的 Hilbert 空间	1.2.4
$\mathscr{R}(z_T)$	z_T 的再生核表示空间	1.2.4
$z(F)$	再生核表示空间的元 F 在 $\mathscr{H}(z_T)$ 中的逆再生表示	1.2.6
$S_z(F, G)$	再生核表示空间的元 F 与 G 的内积(阵)	1.2.6
$S_z(F)$	再生核表示空间的元 F 的平方范数(阵), $S_z(F)\triangleq S_z(F, F)$	1.2.6
$\mathscr{H}^+(z_T)$	$\triangleq \mathscr{H}(z_T)+C$	2.0.1
$\pi(y\|z_T)$	随机向量 y 基于 z_T 的线性最小方差预报	2.1.1
$\hat{\theta}$, $\hat{y}$	回归系数 θ 及随机向量 y 的 GM估计	2.2.1, 2.3.1
$\hat{\hat{\theta}}$, $\hat{\hat{y}}$	回归系数 θ 及随机向量 y 的带线性约束的GM 估计	2.2.3, 2.3.1
$\mu_{\theta\theta}$	回归系数 θ 的先验均值	2.2.5
$\Gamma_{\theta\theta}$	回归系数 θ 的先验方差(阵)	2.2.5
$\check{\theta}$, $\check{y}$	回归系数 θ 及随机向量 y 的 LB 估计	2.2.5, 2.3.4
r_x	x 的信噪比	2.4.1
$P\ll P^{\triangle}$	表示概率测度 P 对 $P^{\triangle}$ 绝对连续	2.5.1
$P\perp P^{\triangle}$	表示概率测度 P 与 $P^{\triangle}$ 相互奇异(正交)	2.5.1
$\frac{dP}{dP^{\triangle}}$	P 对 $P^{\triangle}$ 的 Radon-Nikodym 导数	2.5.1
z_N	二阶随机序列	3.1.1
N_t	不超过 t 的自然数全体	3.1.1
$\mathscr{H}(z, t)$, $\mathscr{H}_t$	$\triangleq \mathscr{H}(z_{N_t})$或 $\mathscr{H}(z_{[0,t]})$	3.1.1, 6.1.1
$\mathscr{R}(z, t)$, $\mathscr{R}_t$	$\triangleq \mathscr{R}(z_{N_t})$或 $\mathscr{R}(z_{[0,t]})$	3.1.1, 6.1.1
$S(F, G)_t$	局限于 N_t 或$[0, t]$上的再生核空间内积阵	3.1.1, 6.1.1
$z(F)_t$	局限于 N_t 或$[0, t]$上的逆再生表示	3.1.1, 6.1.1

e_N	新息序列	3.1.2
Φ_{ts}	转移阵	3.2.1
$\Phi_t^{(Q)}$	由转移阵 Φ 定义的局限于 N_{t-q} 上的矩阵值函数, $(\Phi_t^{(Q)})_s \triangleq \Phi_{t,s+Q}$	3.2.1
$\hat{x}_{\tau\|t}$	x_τ 基于量测 z_{N_t} 或 $z_{[0,t]}$ 的线性最小方差预报	4.0.1, 7.2.4
$P_{\tau\|t}$	$\hat{x}_{\tau\|t}$ 预报 x_τ 的误差方差阵	4.0.1, 7.2.4
$\breve{x}_{\tau\|t}$	x_τ 基于量测 z_{N_t} 或 $z_{[0,t]}$ 的 GM 估计	4.2.1
$\Pi_{\tau\|t}$	$\breve{x}_{\tau\|t}$ 估计 x_τ 的误差方差阵	4.2.1, 7.3.1
$\check{x}_{\tau\|t}$	x_τ 基于量测 z_{N_t} 或 $z_{[0,t]}$ 的 LB 估计	4.2.1
$Q_{\tau\|t}$	$\check{x}_{\tau\|t}$ 估计 x_τ 的平均误差阵	4.2.1, 7.3.1
$(T, \mathscr{T})$	可测空间	5.1.1
1_S	集 S 的示性函数	5.1.2
$(T, \mathscr{T}, \mu)$	阵测度空间	5.1.3
$\mathrm{tr}\,\mu$	阵测度 μ 的迹测度	5.1.4
$\dfrac{d\mu}{d\,\mathrm{tr}\,\mu}$	阵测度对其迹测度的 RN 导数	5.1.4
$\int_T F_t d\mu G_t^*$	矩阵值函数对阵测度的积分	5.1.5
$L_2(\mu)^k$	对 μ 平方可积的 $C^{k\times m}$ 值函数全体	5.1.5
$\sigma L_2(\mu)^k$	对 μ 为 σ 平方可积的 $C^{k\times m}$ 值函数全体	5.1.5
$\int_T H_t d\mu\, z_t$	二阶可积过程 z_T 对阵测度 μ 的随机积分	5.2.3
$w_{\mathscr{S}}$	以集类 $\mathscr{S}$ 为指标集的正交随机测度	5.3.1
$\int_T H_t dw$	平方可积函数对正交随机测度的积分	5.3.4
$(G\cdot w)_{\mathscr{S}}$	由 G 和 $w_{\mathscr{S}}$ 导出的正交随机测度	5.3.7
$L_2(\Phi_T, \nu)$	由函数族 Φ_T 张成的 $L_2(\nu)$ 的闭子空间	5.5.2
$z_{[0,\infty)}$	二阶随机过程	6.0.1
$t'\uparrow\uparrow t$	t' 严格上升趋于 t	6.0.1
$t'\downarrow\downarrow t$	t' 严格下降趋于 t	6.0.1

$H.M$	H对增函数M的不定积分	6.2.3
$\frac{dF}{dM}$	不定积分 F 对 M 的 RN 导数	6.2.4
$\varepsilon_{[0,\infty)}$	(线性)新息过程	6.3.1
Φ	A 的转移阵(函数)	7.1.4
$\hat{x}_t$	x_t 的滤波, $\hat{x}_t \triangleq \hat{x}_{t\mid t}$	7.2.4
P_t	滤波的误差方差阵, $P_t \triangleq P_{t\mid t}$	7.2.4
u_{N}, $u_{[0,\infty)}$	控制策略	8.1.1, 8.2.4
$\beta(u, \tau)$	平均平方损失函数	8.1.3, 8.2.3
$u^0_{N_\tau}$, $u^0_{[0,\tau]}$	N_τ 或[0, τ]上的最优线性控制	8.1.3, 8.2.3

目 录

CONTENTS

第 1 章

二阶随机过程及再生核表示空间

§1.1 记号约定

为了行文的简洁，特对本书采用的某些一般记号作如下约定，以后不再说明. 另一些记号在首次使用时将陆续给出说明.

记号 1.1.1 T 表示任意非空指标集，其元用 t、s 等小写字母记之.

N 表示正整数全体.

C 根据需要，既可理解为复数全体，也可理解为实数全体，这并不影响本书全部内容的叙述与证明.

设 $\mathscr{X}$ 为一集. $\mathscr{X}^{k\times l}$ 表示以 $\mathscr{X}$ 的元为元的 $k\times l$ 矩阵全体. $\mathscr{X}^k \triangleq \mathscr{X}^{k\times 1}$，即以 $\mathscr{X}$ 的元为分量的 k 维列向量全体. 这里及今后，记号“$\triangleq$”表示“定义为”或“记为”.

设 $c\in C^k$, $C\in C^{k\times l}$, c^*、C^* 分别表示向量 c 和矩阵 C 的共轭转置(在实数情形就是转置). $\|c\|\triangleq\sqrt{c^*c}$, $\|C\|\triangleq\sqrt{\mathrm{tr}C^*C}$, 其中 tr 表方阵的迹. 以此为范数，$C^k$ 与 $C^{k\times l}$ 分别构成 k 维与 kl 维(复或实的)欧氏空间.

设 $\Gamma\in C^{k\times k}$ 为一 k 阶 Hermite 阵(在实数情形就是对称阵). Γ^- 表示 Γ 的弱逆，即满足 $\Gamma\Gamma^-\Gamma=\Gamma$ 及 $\Gamma^-\Gamma\Gamma^-=\Gamma^-$ 这两个条件的任一 k 阶 Hermite 阵(相应地，实对称阵). 易知 Γ^- 一定存在，但当 Γ 降秩时不唯一. 又不等式 $\Gamma_1\leqslant\Gamma_2$ 表示对任意 $c\in C^k$ 有 $c^*\Gamma_1c\leqslant c^*\Gamma_2c$ 成立，即 $\Gamma_2-\Gamma_1$ 为半正定阵，也可表示为 Γ_2-

$\Gamma_1\geqslant 0$.

设$(\Omega, \mathscr{F}, P)$为一概率空间，其上的二阶矩有穷的 C 值随机变量全体记作 $\mathscr{L}_2\triangleq\mathscr{L}_2(\Omega, \mathscr{F}, P)$，其中均值为 0 的随机变量全体记作 $\mathscr{L}_{20}$. *对任意 $\boldsymbol{x}\in\mathscr{L}_2^k$, $\boldsymbol{y}\in\mathscr{L}_2^l$, 记 $R(\boldsymbol{x}, \boldsymbol{y})\triangleq E(\boldsymbol{x}-E\boldsymbol{x})\cdot(\boldsymbol{y}-E\boldsymbol{y})^*(\in C^{k\times l})$, $R(\boldsymbol{x})\triangleq R(\boldsymbol{x}, \boldsymbol{x})(\in C^{k\times k})$; $R^+(\boldsymbol{x}, \boldsymbol{y})\triangleq E\boldsymbol{x}\boldsymbol{y}^*$, $R^+(\boldsymbol{x})\triangleq R^+(\boldsymbol{x}, \boldsymbol{x})$. 熟知，如果把 a.s.$(P)$ 相等的随机变量视为同一，则 $\mathscr{L}_2$ 构成以 R^+ 为内积的 Hilbert 空间，$\mathscr{L}_{20}$ 为其闭子空间. 在 $\mathscr{L}_{20}$ 上，$R^+=R$. 有时为了明确表出其内积定义，这两个 Hilbert 空间也分别记作$(\mathscr{L}_2, R^+)$, $(\mathscr{L}_{20}, R)$. 今后，随机变量的相等都是指 a.s. 相等，并不再标明 a.s.

下面是两个以后常用的简单结果.

引理 1.1.2　设 $\boldsymbol{x}\in\mathscr{L}_2^k$, $\boldsymbol{y}\in\mathscr{L}_2^l$, 则

$$R(\boldsymbol{x}, \boldsymbol{y})=R(\boldsymbol{x}, \boldsymbol{y})R(\boldsymbol{y})^-R(\boldsymbol{y})$$

$$\|R(\boldsymbol{x}, \boldsymbol{y})\|^2\leqslant \operatorname{tr}R(\boldsymbol{x})\operatorname{tr}R(\boldsymbol{y}).$$

证　首先易知，

$$\begin{aligned}R(\boldsymbol{y}-R(\boldsymbol{y})R(\boldsymbol{y})^-\boldsymbol{y})&=R(\boldsymbol{y})+R(\boldsymbol{y})R(\boldsymbol{y})^-R(\boldsymbol{y})R(\boldsymbol{y})^-R(\boldsymbol{y})\\&\quad-R(\boldsymbol{y})R(\boldsymbol{y})^-R(\boldsymbol{y})-R(\boldsymbol{y})R(\boldsymbol{y})^-R(\boldsymbol{y})\\&=2R(\boldsymbol{y})-2R(\boldsymbol{y})=0,\end{aligned}$$

故 $\boldsymbol{y}=R(\boldsymbol{y})R(\boldsymbol{y})^-\boldsymbol{y}+\boldsymbol{c}$($\boldsymbol{c}$ 为常向量)，于是

$$R(\boldsymbol{x}, \boldsymbol{y})=R(\boldsymbol{x}, R(\boldsymbol{y})R(\boldsymbol{y})^-\boldsymbol{y})=R(\boldsymbol{x}, \boldsymbol{y})R(\boldsymbol{y})^-R(\boldsymbol{y}).$$

其次，分别记 $\boldsymbol{x}$, $\boldsymbol{y}$ 的分量为 $x^{(1)}, \cdots, x^{(k)}$; $y^{(1)}, \cdots, y^{(l)}$. 则由关于内积的 Schwarz 不等式，有

$$\begin{aligned}\|R(\boldsymbol{x}, \boldsymbol{y})\|^2&=\operatorname{tr}[R(\boldsymbol{x}, \boldsymbol{y})R(\boldsymbol{y}, \boldsymbol{x})]=\sum_{i=1}^k\sum_{j=1}^l|R(x^{(i)}, y^{(j)})|^2\\&\leqslant\sum_{i=1}^k R(x^{(i)})\sum_{j=1}^l R(y^{(j)})=\operatorname{tr}R(\boldsymbol{x})\operatorname{tr}R(\boldsymbol{y}).\end{aligned}$$ ∎

§1.2　正则过程的生成空间与再生核表示空间

本节要讨论的基本问题，是如何将观测一个任意指标集上的随机过程所能获得的全部“线性统计量”与该指标集上的一类函数

构成一一同构对应. 这种对应就称为"表示". 它是本书其余一切内容的基础. 最先详尽讨论这一问题的是 J. Hájek[1].

定义 1.2.1 设 $\boldsymbol{T}$ 为一非空指标集. 若任一 $t\in\boldsymbol{T}$, 有一 $z_t\in\mathscr{L}_2^m$ 与之对应, 则称 $z_T\triangleq(z_t,\ t\in\boldsymbol{T})$ 为 $\boldsymbol{T}$ 上的一个 m 维二阶随机过程, 或简称为 m 维二阶过程.

设 z_T 为一 m 维二阶过程. 记

$$\mathscr{H}_0(z_T)\triangleq\left\{\sum_{i=1}^{n}c_i^*z_{t_i}:\ n\in\boldsymbol{N},\ c_i\in\boldsymbol{C}^m,\ t_i\in\boldsymbol{T}\right\},\quad(1.2.1)$$

称为由 z_T 张成的线性空间. 又对任一 $x\in\mathscr{L}_2^k$, 令

$$Z_t^x\triangleq R(x,\ z_t),\ t\in\boldsymbol{T},\qquad(1.2.2)$$

则 Z^x 为 $\boldsymbol{T}$ 上的 $\boldsymbol{C}^{k\times m}$ 值函数. 注意: 这里及今后, 对 $x\in\mathscr{L}_2^k$, 我们总是用与过程符号相应的大写字母添加右上角标 x(例如, 对过程 z_T、w_T、y_T 等, 分别用 Z^x, W^x, Y^x 等)来表记这个与该过程及 x 相联系的矩阵值函数.

显然, 对任意 $x\in\mathscr{L}_2^k$, $y\in\mathscr{L}_2^l$, $C\in\boldsymbol{C}^{p\times k}$, $D\in\boldsymbol{C}^{p\times l}$, 总有 $Cx+Dy\in\mathscr{L}_2^p$, 且有

$$Z^{Cx+Dy}=CZ^x+DZ^y\qquad(1.2.3)$$

成立.

又记

$$\begin{aligned}\mathscr{R}_0(z_T)&\triangleq\{Z^x:x\in\mathscr{H}_0(z_T)\}\\&=\left\{\sum_{i=1}^{n}c_i^*Z^{z_{t_i}}:\ n\in\boldsymbol{N},\ c_i\in\boldsymbol{C}^m,\ t_i\in\boldsymbol{T}\right\}.\quad(1.2.4)\end{aligned}$$

这是一个由 $\boldsymbol{T}$ 上的 $\boldsymbol{C}^{1\times m}$ 值函数组成的线性空间.

定义 1.2.2 设 z_T 为一 m 维二阶过程. 若存在 $\alpha\geqslant0$, 使对任意 $x\in\mathscr{H}_0(z_T)$ 有

$$|Ex|^2\leqslant\alpha R(x),$$

则称 z_T 为 m 维二阶正则过程, 或简称为 m 维正则过程*).

显然, 均值为 0 的二阶过程是正则过程的特例.

引理 1.2.3 设 z_T 为一 m 维正则过程, 则 $(\mathscr{H}_0(z_T),\ R)$ 为

---

*) 正则过程在本书中是一个很基本的概念, 它的重要意义通过后面的定理 1.2.11 及定义 1.2.13 的说明, 尤其是通过第二章的讨论, 可逐渐明白.

一内积空间，且在 $\mathscr{H}_0(z_T)$ 上，由 R 与 R^+ 导出的范数等价.

证 为了证明 R 是线性空间 $\mathscr{H}_0(z_T)$ 的内积，只须证：对 $x\in\mathscr{H}_0(z_T)$，$R(x)=0$ 蕴含 $x=0$. 事实上，由定义1.2.2，对任意 $x\in\mathscr{H}_0(z_T)$ 有

$$R(x)\leqslant R^+(x)=E|x|^2=R(x)+|Ex|^2\leqslant(1+\alpha)R(x).$$

所以由 $R(x)=0$ 可得 $R^+(x)=0$，从而 $x=0$；且知由 R 与 R^+ 在 $\mathscr{H}_0(z_T)$ 上导出的范数为等价. ∎

定义1.2.4 设 z_T 为一 m 维正则过程，令 $(\mathscr{H}(z_T),\ R)$ 为内积空间 $(\mathscr{H}_0(z_T),\ R)$ 的完备化，称为由 z_T 生成的 Hilbert 空间. 由引理1.2.3，$\mathscr{H}(z_T)$ 为 $\mathscr{L}_2$ 的闭子空间.

又记

$$\mathscr{R}(z_T)\triangleq\{Z^x;\ x\in\mathscr{H}(z_T)\},\tag{1.2.5}$$

显然有

$$\mathscr{R}(z_T)^k=\{Z^x;\ x\in\mathscr{H}(z_T)^k\},$$

$\mathscr{R}(z_T)^k$ 的元是 T 上的 $C^{k\times m}$ 值函数，本书常用 F、G 等大写字母标记；有时也特别用 f、g 等小写字母表示 $\mathscr{R}(z_T)$ 的元，它们是 T 上的 $C^{1\times m}$ 值(行向量)函数.

上述定义中 $\mathscr{H}(z_T)$ 中的元就是由 z_T 构成的线性统计量. 从数学上来看，当指标集 T 为无限集时，将 $\mathscr{H}_0(z_T)$ 完备化为 $\mathscr{H}(z_T)$ 是十分必要的，否则，很多估计的存在性就成了问题.

定理1.2.5 设 z_T 为一 m 维正则过程，则映射 $x\mapsto Z^x$ 是 $\mathscr{H}(z_T)^k$ 映上 $\mathscr{R}(z_T)^k$ 的线性一一映射.

证 映射的线性性质已由(1.2.3)式给出，剩下的是证明一一性. 不失普遍性，可设 $k=1$. 设 $x\in\mathscr{H}(z_T)$ 并且 $Z^x=0$，即 $R(x,z_t)=0$，$t\in T$. 则对任一 $y=\sum_{i=1}^n c_i^* z_{t_i}\in\mathscr{H}_0(z_T)$ 有

$$R(x,\ y)=\sum_{i=1}^n R(x,\ z_{t_i})c_i=0.$$

因为 $x\in\mathscr{H}(z_T)$，存在 $(x_n)\subset\mathscr{H}_0(z_T)$，使当 $n\to\infty$ 时，$R(x-x_n)\to0$，且由上式知 $R(x,\ x_n)=0$. 于是

$$R(x)=R(x,\ x)-R(x,\ x_n)=R(x,\ x-x_n)$$

$$\leqslant\sqrt{R(x)}\sqrt{R(x-x_n)}\to 0.$$

所以 $R(x)=0$. 由过程的正则性知 $R^+(x)=0$, 因而 $x=0$. ▮

本节开头提到的一一对应，就是本定理中指明的 $\mathscr{H}(z_T)$ 和 $\mathscr{R}(z_T)$ 间的一一线性映射. 这个对应关系的重要性可在第二章中看到.

定义 1.2.6 设 z_T 为一 m 维正则过程. 由定理 1.2.5, 对任一 $F\in\mathscr{R}(z_T)^k$, 必存在唯一的 $z(F)\in\mathscr{H}(z_T)^k$, 使 $F=Z^{z(F)}$, 即有 $F_t=R(z(F),z_t)$, $t\in\boldsymbol{T}$. 换句话说，对于空间 $\mathscr{H}(z_T)^k$ 与 $\mathscr{R}(z_T)^k$, 映射 $F\mapsto z(F)$ 是 $x\mapsto Z^x$ 的逆映射. 注意，这里及以后，我们总是用与过程符号相同的小写字母(例如，对过程 z_T、w_T、y_T 等，分别用 $z(\cdot)$、$w(\cdot)$、$y(\cdot)$ 等)来表记这个逆映射. 显然它也是线性的，即对任意 $F\in\mathscr{R}(z_T)^k$, $G\in\mathscr{R}(z_T)^l$, $C\in\boldsymbol{C}^{p\times k}$, $D\in\boldsymbol{C}^{p\times l}$, 有 $CF+DG\in\mathscr{R}(z_T)^p$, 且有下式成立:

$$z(CF+DG)=Cz(F)+Dz(G).\tag{1.2.6}$$

又令

$$S_z(F,G)\triangleq R(z(F),z(G)),\ S_z(F)\triangleq S_z(F,F)=R(z(F)).\tag{1.2.7}$$

下一定理概括了空间 $(\mathscr{R}(z_T),S_z)$ 的构造与基本性质.

定理 1.2.7 设 z_T 为一 m 维正则过程，$\mathscr{R}_0(z_T)$、$\mathscr{R}(z_T)$ 与 S_z 分别如(1.2.4)、(1.2.5)、(1.2.7)所定义. 记 z_T 的协方差函数为

$$\Gamma(t,s)\triangleq R(z_t,z_s),\ (t、s\in\boldsymbol{T}),\tag{1.2.8}$$

且对任一 $t\in\boldsymbol{T}$, 用 Γ^t 记如下规定的 $m\times m$ 阵函数:

$$\Gamma^t_s\triangleq\Gamma(t,s),\ (s\in\boldsymbol{T}),\tag{1.2.9}$$

则有

i) $\mathscr{R}_0(z_T)=\left\{\sum_{i=1}^n c_i^*\Gamma^{t_i}:\ n\in\boldsymbol{N},\ c_i\in\boldsymbol{C}^m,\ t_i\in\boldsymbol{T}\right\}.$

ii) $(\mathscr{R}_0(z_T),S_z)$ 为一内积空间，且

$$S_z\left(\sum_i c_i^*\Gamma^{t_i},\ \sum_j d_j^*\Gamma^{s_j}\right)=\sum_{i,j}c_i^*\Gamma(t_i,s_j)d_j.$$

iii) 若 (f_n) 为 $(\mathscr{R}_0(z_T),S_z)$ 中的基本列，即当 n、$l\to\infty$ 时，

$S_s(f_n-f_l)\to 0$，则对任意 $t\in T$，(f_{nt})是$(C^{1\times m}, \| \ \|)$中的收敛列.

iv) $\mathscr{R}(z_T)=\overline{(\mathscr{R}_0(z_T), S_s)}$(即 $\mathscr{R}_0(z_T)$ 在 S_s 下的完备化)，$(\mathscr{R}(z_T), S_s)$为一Hilbert 空间.

v) 对任意 $F\in\mathscr{R}(z_T)^k$，有

$$F_t=S_s(F, \Gamma^t), \ (t\in T) \tag{1.2.10}$$

成立.

vi) 若 $F\in\mathscr{R}(z_T)^k$, $G\in\mathscr{R}(z_T)^l$, 则

$$\|S_s(F, G)\|^2\leqslant \operatorname{tr} S_s(F)\operatorname{tr} S_s(G). \tag{1.2.11}$$

vii) 若 $(F_n)\subset\mathscr{R}(z_T)^k$, $F\in\mathscr{R}(z_T)^k$, 且当 $n\to\infty$ 时，$S_s(F_n-F)\to 0$，则对任意 $t\in T$ 有 $\|F_{nt}-F_t\|\to 0$. 即依范数 S_s 收敛蕴含点点收敛；若 Γ 为有界的，则更蕴含一致收敛.

证 i) 由定义，阵函数 $\Gamma^t=Z^{z_t}$. 故从(1.2.4)即得所证.

ii) 由映射 $F\mapsto z(F)$的线性一一性质，(1.2.7)定义的S_s显然是 $\mathscr{R}(z_T)$上的内积，且从 $z(\Gamma^t)=z_t$ 得

$$\begin{aligned}
&S_s(\sum_i c_i^*\Gamma^{t_i}, \sum_j d_j^*\Gamma^{s_j})\\
&=R(\sum_i c_i^* z_{t_i}, \sum_j d_j^* z_{s_j})\\
&=\sum_{i,j} c_i^* R(z_{t_i}, z_{s_j})d_j=\sum_{i,j} c_i^*\Gamma(t_i, s_j)d_j.
\end{aligned}$$

iii) 由引理1.1.2, 对任意 $t\in T$ 有

$$\begin{aligned}
\|f_{nt}-f_{lt}\|^2&=\|R(z(f_n-f_l), z_t)\|^2\\
&\leqslant R(z(f_n-f_l))\operatorname{tr} R(z_t)\\
&=S_s(f_n-f_l)\operatorname{tr}\Gamma(t, t)\to 0, \ n, l\to\infty.
\end{aligned}$$

再由空间$(C^{1\times m}, \| \ \|)$的完备性即知(f_{nt})是其中的收敛列.

iv) 因为 $\mathscr{H}(z_T)=\overline{(\mathscr{H}_0(z_T), R)}$，而 $\mathscr{R}_0(z_T)$与 $\mathscr{R}(z_T)$在映射 $f\mapsto z(f)$下分别和 $\mathscr{H}_0(z_T)$与 $\mathscr{H}(z_T)$是保内积线性同构的，所以结论成立.

v) 设 $F\in\mathscr{R}(z_T)^k$, 则由(1.2.7)，对任意 $t\in T$ 有

$$F_t=R(z(F), z_t)=S_s(F, \Gamma^t).$$

vi) 由(1.2.7)和引理1.1.2.

vii) 仿 iii) 的证明或直接由 vi), 对任意 $t\in \boldsymbol{T}$ 有

$$\|F_{nt}-F_t\|^2\leqslant \operatorname{tr}\Gamma(t,\ t)\operatorname{tr}S_s(F_n-F).\ \blacksquare$$

注记1.2.8

i) 定理1.2.7 表明, 虽然我们是从一个m维二阶正则过程出发来定义空间$(\mathscr{R}(z_T),\ S_s)$的, 但后者实质上只依赖于z_T的协方差阵函数Γ, 而与 z_T 的其他统计性质(例如, z_T 的均值及其分布的高于二阶矩的特征)无关. 只要给定Γ, 经过步骤 i)到 iv), 就可以唯一地确定出这个 Hilbert 空间 (其中的 iii) 说明 $\mathscr{R}_0(z_T)$的完备化元仍是 $\boldsymbol{T}$ 上的 $C^{1\times m}$ 值函数).

ii) (1.2.10) 式表明, $\mathscr{R}(z_T)^k$ 的任一元($\boldsymbol{T}$ 上的 $C^{k\times m}$ 值函数)可以通过该元与一确定的核函数 ($\boldsymbol{T}^2$ 上的阵函数)Γ 的截口作内积"再生"出来. 因此通常称$(\mathscr{R}(z_T),\ S_s)$ 为以 Γ 为再生核的 Hilbert 空间,或与 z_T 的 协方差 R 相应的再生核空间.

下一个定理给出了 $\mathscr{R}(z_T)$中元的一个重要刻划.

定理1.2.9 设 $\Gamma(t,\ s)$, t、$s\in \boldsymbol{T}$ 为m维二阶过程 z_T 的协方差阵函数, $(\mathscr{R}(z_T),\ S_s)$为以 Γ 为再生核的 Hilbert 空间. 若 F 为一 $\boldsymbol{T}$ 上的 $C^{k\times m}$ 值函数, 则 $F\in\mathscr{R}(z_T)^k$ 的必要与充分条件是: 存在 $\alpha>0$, 使对任意有限集 $\{t_i\}\subset \boldsymbol{T}$, $\{c_i\}\subset C^m$ 有

$$\|\sum_i F_{t_i}c_i\|^2\leqslant\alpha\sum_{i,j}c_i^*\Gamma(t_i,\ t_j)c_j=\alpha R(\sum_i c_i^* z_{t_i}).\quad(1.2.12)$$

证 必要性: 设 $F\in\mathscr{R}(z_T)^k$, 则由(1.2.10)和(1.2.11),

$$\begin{aligned}\|\sum_i F_{t_i}c_i\|^2&=\|\sum_i S_s(F,\ \Gamma^{t_i})c_i\|^2=\|S_s(F,\ \sum_i c_i^*\Gamma^{t_i})\|^2\\&\leqslant \operatorname{tr}S_s(F)S_s(\sum_i c_i^*\Gamma^{t_i})\\&=\operatorname{tr}S_s(F)\sum_{i,j}c_i^*\Gamma(t_i,\ t_j)c_j.\end{aligned}$$

令 $\alpha=\operatorname{tr}S_s(F)$, 即得所证.

充分性: 设 F 使不等式(1.2.12)成立. 不妨设 z_T 的均值为0, 因而为正则. 否则, 可考虑减去均值函数的过程. 令

$$\Phi(x)\triangleq(\sum_i F_{t_i}c_i)^*,\quad x=\sum_i c_i^* z_{t_i}\in\mathscr{H}_0(z_T).$$

则 $\Phi(x)$ 对 x 是线性的，而 (1.2.12) 表明 Φ 的定义是确定的且 $\|\Phi(x)\|^2\leqslant \alpha R(x)$，即 Φ 是内积空间 $(\mathscr{H}_0(z_T), R)$ 上的 k 维有界线性泛函. 通过取极限容易将 Φ 唯一地扩充为 Hilbert 空间 $(\mathscr{H}(z_T), R)$ 上的 k 维有界线性泛函. 再由 Hilbert 空间的自伴性 (Riesz 表现定理)，存在唯一的 $y\in\mathscr{H}(z_T)^k$，使对任意 $x\in\mathscr{H}(z_T)$，有 $\Phi(x)=R(x, y)$ 成立. 特别是，对任意 $c\in C^m$，$t\in T$ 有

$$(F_t c)^* = \Phi(c^* z_t) = R(c^* z_t,\ y) = c^* R(z_t,\ y) = (R(y,\ z_t)c)^*,$$

于是得

$$F_t = R(y,\ z_t),\ (t\in T).$$

这就证明了 $F\in\mathscr{R}(z_T)^k$，且 $z(F)=y$. ▌

系 1.2.10　对任一 $y\in\mathscr{L}_2^k$ 恒有 $Z^y\in\mathscr{R}(z_T)^k$.

证　由引理 1.1.2，对任意有限集 $\{t_i\}\subset T$，$\{c_i\}\subset C^m$，

$$\Big\|\sum_i Z^y_{t_i} c_i\Big\|^2 = \Big\|\sum_i R(y,\ z_{t_i})c_i\Big\|^2 = \Big\|R\Big(y,\ \sum_i c_i^* z_{t_i}\Big)\Big\|$$
$$\leqslant \operatorname{tr} R(y) R\Big(\sum_i c_i^* z_{t_i}\Big).$$

由定理 1.2.9 即得所证. ▌

下面给出正则过程的一个重要刻划.

定理 1.2.11　设 z_T 为一 m 维二阶过程. 令

$$f_t \triangleq (Ez_t)^* = Ez_t^*,\ (t\in T),$$

则 z_T 为正则过程的必要充分条件是 $f\in\mathscr{R}(z_T)$.

证　由定义 1.2.2，z_T 为正则过程，当且仅当存在 $\alpha\geqslant 0$，使任意 $x=\sum_{i=1}^n c_i^* z_{t_i}\in\mathscr{H}_0(z_T)$ 有

$$\Big|\sum_{i=1}^n f_{t_i} c_i\Big|^2 = \Big|\sum_{i=1}^n c_i^*\, Ez_{t_i}\Big|^2 = |Ex|^2 \leqslant \alpha R(x)$$
$$= \alpha R\Big(\sum_{i=1}^n c_i^* z_{t_i}\Big).$$

而由定理 1.2.9，这正等价于 $f\in\mathscr{R}(z_T)$. ▌

下面是第二章常用的一个关键性定理.

定理 1.2.12　设 z_T 为一 m 维正则过程，$f_t\triangleq Ez_t^*$　$(t\in T)$，则对任一 $x\in\mathscr{H}(z_T)^k$ 有

$$Ex = R(x,\ z(f)) = S_z(Z^x,\ f), \qquad (1.2.13)$$

或者等价地，对任一 $F\in\mathscr{R}(z_T)^k$ 有

$$Ez(F)=S_s(F,\ f). \tag{1.2.14}$$

证　首先设 $x=\sum_i c_i^* z_{t_i}\in\mathscr{H}_0(z_T)$. 则由定理 1.2.11 得

$$Ex=\sum_i c_i^* f_{t_i}^*=(\sum_i f_{t_i}c_i)^*=(\sum_i R(z(f),\ z_{t_i})c_i)^*$$
$$=(R(z(f),\ \sum_i c_i^* z_{t_i}))^*=R(x,\ z(f)).$$

其次，在上式两端对 $\mathscr{H}_0(z_T)$ 中的基本列取极限，注意两端都是 $(\mathscr{H}_0(z_T),\ R)$ 上的连续泛函，即知(1.2.13)对任意 $x\in\mathscr{H}(z_T)$ 成立，从而对任意 $x\in\mathscr{H}(z_T)^k$ 也成立. ▌

定义 1.2.13　设 z_T 为一 m 维正则过程. 则 $(\mathscr{R}(z_T),\ S_s)$ 称为 z_T 的再生核表示空间. $\mathscr{R}(z_T)^k$ 到 $\mathscr{H}(z_T)^k$ 上的线性保内积映射 $F\mapsto z(F)$ 称为逆再生表示.

注意：对非正则的二阶过程，$(\mathscr{H}_0(z_T),\ R)$ 一般不再是内积空间. 当然我们可以考虑内积空间 $(\mathscr{H}_0(z_T),\ R^+)$，并取它在空间 $\mathscr{L}_2$ 中的闭包作为 z_T 的生成空间. 但与 R^+ 相应的再生核空间及逆再生表示都与过程的均值有关，而在线性统计问题中，过程的均值(或其中未知参数)常常是待估计的，故这样的逆表示一般不能构成"线性统计量". 克服非正则性困难的一个办法，是设法对 z_T 虚拟另一确定的协方差函数 $R^\triangle$ 而其均值不变，使它在 $R^\triangle$ 下成为正则过程. 这时 $(\mathscr{H}_0(z_T),\ R^\triangle)$ 成为内积空间，而它的完备化 $(\mathscr{H}^\triangleright(z_T),\ R^\triangle)$ 就可作为线性统计量组成的 Hilbert 空间. 这个办法还可用来处理 z_T 的协方差 R 未知的情形(详情见第二章注记 2.2.11). 下一节讨论的内容正是与解决这个问题密切关联的.

§1.3　协方差改变的情形

定义 1.3.1　设 $(\Omega,\mathscr{F})$ 为一可测空间，P、$P^\triangle$ 是其上的两个概率测度，其相应的期望算子分别记为 E 与 $E^\triangle$. 仿前记

$$R^\triangle(x,\ y)\triangleq E^\triangle(x-E^\triangle x)(y-E^\triangle y)^*,\quad R^\triangle(x)\triangleq R^\triangle(x,\ x).$$

若 $z_T\triangleq(z_t,\ t\in T)$ 关于 P、$P^\triangle$ 都是 m 维二阶过程，且存在 $\beta>0$，

使对任意 $x\in\mathscr{H}_0(z_T)$ 有 $R(x)\leqslant\beta R^{\Delta}(x)$. 则称 z_T 的协方差 R 被 R^{Δ} 控制，或称 R^{Δ} 控制 R.

换句话说，若记 z_T 的两个协方差阵函数分别为

$\Gamma(t, s)\triangleq R(z_t, z_s)$, $\Gamma^{\Delta}(t, s)\triangleq R^{\Delta}(z_t, z_s)$, $(t、s\in\boldsymbol{T})$,

则 z_T 的协方差 R 被 R^{Δ} 控制，是指存在 $\beta>0$, 使对任意有限集 $\{t_i\}\subset\boldsymbol{T}$, $\{c_i\}\subset\boldsymbol{C}^m$, 有

$$\sum_{i,j}c_i^*[\beta\Gamma^{\Delta}(t_i, t_j)-\Gamma(t_i, t_j)]c_j\geqslant 0$$

成立.

定理1.3.2 设 z_T 关于 P、P^{Δ} 都是 m 维二阶过程，且关于 P^{Δ} 是正则过程. 记后者生成的 Hilbert 空间为 $(\mathscr{H}^{\Delta}(z_T), R^{\Delta})$. 则 R 被 R^{Δ} 控制的必要充分条件是：存在 $\mathscr{H}^{\Delta}(z_T)\to\mathscr{H}^{\Delta}(z_T)$ 的线性有界正(自伴)算子 $\boldsymbol{B}$, 使对任意 x、$y\in\mathscr{H}_0(z_T)$ 有 $R(x, y)=R^{\Delta}(x, \boldsymbol{B}y)$. 算子 R 的范数为

$$\|\boldsymbol{B}\|=\sup_{x\in\mathscr{H}_0(z_T)}R(x)/R^{\Delta}(x)<\infty.$$

证　充分性：若所述的算子 $\boldsymbol{B}$ 存在，则对任意 $x\in\mathscr{H}_0(z_T)$,

$$R(x)=R^{\Delta}(x, \boldsymbol{B}x)\leqslant\sqrt{R^{\Delta}(x)}\sqrt{R^{\Delta}(\boldsymbol{B}x)}\leqslant\|\boldsymbol{B}\|R^{\Delta}(x),$$

故 R 被 R^{Δ} 控制.

必要性：设存在 $\beta>0$, 使对任意 $x\in\mathscr{H}_0(z_T)$ 有 $R(x)\leqslant\beta R^{\Delta}(x)$. 则显然对任意 $x\in\mathscr{H}^{\Delta}(z_T)$, 此不等式也成立(注意：这里的 x 实际上意味若与 x a.s. P^{Δ} 相等的任一随机变量). 于是对任意 x、$y\in\mathscr{H}^{\Delta}(z_T)$ 有

$$|R(x, y)|\leqslant\sqrt{R(x)}\sqrt{R(y)}\leqslant\beta\sqrt{R^{\Delta}(x)}\sqrt{R^{\Delta}(y)},$$

故对任一 $y\in\mathscr{H}^{\Delta}(z_T)$, $R(\cdot, y)$ 是 Hilbert 空间 $(\mathscr{H}^{\Delta}(z_T), R^{\Delta})$ 上的有界线性泛函，因而存在唯一的 $\boldsymbol{B}y\in\mathscr{H}^{\Delta}(z_T)$ 使

$$R(x, y)=R^{\Delta}(x, \boldsymbol{B}y),\ x\in\mathscr{H}^{\Delta}(z_T).$$

将 $\boldsymbol{B}$ 视为 $\mathscr{H}^{\Delta}(z_T)\to\mathscr{H}^{\Delta}(z_T)$ 的算子. 显然它是线性的. 又

$$\sqrt{R^{\Delta}(\boldsymbol{B}y)}=\sup_{x\in\mathscr{H}_0(z_T)}\frac{|R(x, y)|}{\sqrt{R^{\Delta}(x)}}$$

$$\leqslant \sup_{x\in\mathscr{H}_0(z_T)} \frac{\beta\sqrt{R^{\Delta}(x)}\sqrt{R^{\Delta}(y)}}{\sqrt{R^{\Delta}(x)}} = \beta\sqrt{R^{\Delta}(y)},$$

所以 $\boldsymbol{B}$ 是有界的，且 $\|\boldsymbol{B}\|\leqslant\beta$. 又由其中的第一个等式知

$$\sup_{x\in\mathscr{H}_0(z_T)} \frac{R(x)}{R^{\Delta}(x)} \leqslant \sup_{x,y\in\mathscr{H}_0(z_T)} \frac{|R(x, y)|}{\sqrt{R^{\Delta}(x)}\sqrt{R^{\Delta}(y)}}$$

$$= \sup_{y\in\mathscr{H}_0(z_T)} \frac{\sqrt{R^{\Delta}(\boldsymbol{B}y)}}{\sqrt{R^{\Delta}(y)}} \triangleq \|\boldsymbol{B}\|$$

$$\leqslant \inf\{\beta;\ R(x)\leqslant\beta R^{\Delta}(x),\ x\in\mathscr{H}_0(z_T)\}$$

$$= \sup_{x\in\mathscr{H}_0(z_T)} \frac{R(x)}{R^{\Delta}(x)}.$$

最后，由于

$$R^{\Delta}(x, \boldsymbol{B}y) = R(x, y) = R(y, x)^* = R^{\Delta}(y, \boldsymbol{B}x)^*$$

$$= R^{\Delta}(\boldsymbol{B}x, y)\text{（自伴性）},$$

$$R^{\Delta}(x, \boldsymbol{B}x) = R(x) \geqslant 0\text{（半正定性）},$$

所以 $\boldsymbol{B}$ 是正算子. ▌

定理1.3.3 设 m 维二阶过程 z_T 的协方差 R 被 R^{Δ} 控制，分别记其相应的再生核空间为 $(\mathscr{R}(z_T), S_z)$, $(\mathscr{R}^{\Delta}(z_T), S_z^{\Delta})$, 则 $\mathscr{R}(z_T)\subset R^{\Delta}(z_T)$, 且对任意 $F\in\mathscr{R}(z_T)^k$, 有

$$S_z^{\Delta}(F)\leqslant\beta S_z(F)$$

成立. 其中 $$\beta\triangleq\sup_{x\in\mathscr{H}_0(z_T)} R(x)/R^{\Delta}(x).$$

证 首先由注记1.2.8知, $\mathscr{R}(z_T)$, $\mathscr{R}^{\Delta}(z_T)$ 的构造只与 R、R^{Δ} 有关，而无须假定 z_T 的正则性，且由定理1.2.9即知 $\mathscr{R}(z_T)\subset\mathscr{R}^{\Delta}(z_T)$.

其次，不妨设 z_T 关于 P^{Δ} 的均值函数为 0, 因而是正则过程. 于是由定理1.3.2, 存在 $\mathscr{H}^{\Delta}(z_T)\to\mathscr{H}^{\Delta}(z_T)$ 的正线性算子 $\boldsymbol{B}$, 使对任意 x、$y\in\mathscr{H}_0(z_T)$ 有 $R(x, y) = R^{\Delta}(x, \boldsymbol{B}y) = R^{\Delta}(\boldsymbol{B}x, y)$ 成立，特别有 $f_t\triangleq R(x, z_t) = R^{\Delta}(\boldsymbol{B}x, z_t)$, $(t\in\boldsymbol{T})$. 故 $S_z(f) = R(x)$, 而 f 在 $\mathscr{H}^{\Delta}(z_T)$ 中的逆再生表示像为 $z^{\Delta}(f) = \boldsymbol{B}x$. 于是有

$$S_z^{\triangle}(f)=R^{\triangle}(z^{\triangle}(f))=R^{\triangle}(Bx)=\sup_{y\in\mathscr{R}_0(z_T)}|R(y,x)|^2/R^{\triangle}(y)$$
$$\leqslant\sup_{y\in\mathscr{R}_0(z_T)}\beta R^{\triangle}(y)R(x)/R^{\triangle}(y)=\beta R(x)=\beta S_z(f),$$

即 $S_z^{\triangle}(f)\leqslant\beta S_z(f)$ 对任意 $f\in\mathscr{R}_0(z_T)$ 成立. 两边对 $(\mathscr{R}_0(z_T),\ S_z)$ 中的基本列取极限, 即知不等式对任意 $f\in\mathscr{R}(z_T)$ 成立.

最后, 设 $F\in R(z_T)^k$, 则对任意的 $c\in C^k$ 有 $c^*F\in\mathscr{R}(z_T)$. 故有

$$c^*S_z^{\triangle}(F)c=S_z^{\triangle}(c^*F)\leqslant\beta S_z(c^*F)=\beta c^*S_z(F)c,$$

即有 Hermite 阵不等式 $S_z^{\triangle}(F)\leqslant\beta S_z(F)$ 成立. ▌

§1.4 指标集为单元集的情形

如果 z_T 是某个统计问题中被观测的样本, 则在经典统计学(包括时间序列分析)中, 指标集 T 通常都是有限集. 而最简单且基本的情形则是 T 为单元集的情形. 这时, z_T 就是一个 m 维二阶随机向量 z; $R(z)$ 是 z 的协方差阵, 而以 $B(z)$ 为再生核的 Hilbert 空间 $\mathscr{R}(z)$ 则是由一些 m 维行向量组成的 $(\subset C^{1\times m})$.

定理 1.4.1 设 T 为单元集, $z_T=z\in\mathscr{L}_2^m$, 则

$$\mathscr{R}(z)^k=\{F: F=\tilde{F}R(z),\ \tilde{F}\in C^{k\times m}\},$$

即 $\mathscr{R}(z)$ 是由 z 的协方差阵 $R(z)$ 的行向量张成的线性空间. 对任意 $F\in\mathscr{R}(z)^k$, $G\in\mathscr{R}(z)^l$ 有 $S_z(F,G)=FR(z)^-G^*$, $S_z(F)=FR(z)^-F^*$. 又当且仅当 $R(z)$ 为满秩时, $\mathscr{R}(z)^k=C^{k\times m}$, 此时 $R(z)^-=R(z)^{-1}$, 且 z 必是正则的.

一般, 若 z 是正则的, 即 $Ez^*\in\mathscr{R}(z)$, 则 $F\in\mathscr{R}(z)^k$ 的逆再生表示像为 $z(F)=FR(z)^-z$.

证 记 $\Gamma\triangleq R(z)$. 不妨设 z 是正则的, 易知此时(因为有限维内积空间总是完备的)

$$\mathscr{H}(z)^k=\mathscr{H}_0(z)^k=\{\tilde{F}z: \tilde{F}\in C^{k\times m}\},$$

于是有

$$\mathscr{R}(z)^k=\{F\triangleq R(\tilde{F}z,z)=\tilde{F}\Gamma: \tilde{F}\in C^{k\times m}\}.$$

显然，当且仅当 Γ 为满秩时，方程 $F=X\Gamma$ 对任意 $F\in C^{k\times m}$ 有解 $X=F\Gamma^{-1}\in C^{k\times m}$.

设 $F=\tilde{F}\Gamma\in\mathscr{R}(z)^k$，则

$$F=\tilde{F}\Gamma\Gamma^-\Gamma=F\Gamma^-\Gamma=R(F\Gamma^-z,\ z),$$

即得 $z(F)=F\Gamma^-z$，且对任意 $F\in\mathscr{R}(z)^k$，$G\in\mathscr{R}(z)^l$ 有

$$S_z(F,\ G)=R(F\Gamma^-z,\ G\Gamma^-z)=F\Gamma^-\Gamma\Gamma^-G^*=F\Gamma^-G^*.\ \blacksquare$$

注记 1.4.2 在 T 为单元集的情形，z 的协方差 R 被 R^{Δ} 控制的必要充分条件显然是：存在 $\beta>0$，使 z 的相应的协方差阵 $R(z)$ 与 $R^{\Delta}(z)$ 满足 $R(z)\leqslant\beta R^{\Delta}(z)$. 若 $R^{\Delta}(z)=I$（单位阵），则只要取 β 为 $R(z)$ 的最大本征值即可使不等式成立.

又在经典统计学最常见的线性模型中，$R(z)$ 总是假设为满秩阵，甚至是更特殊的 $R(z)=\sigma^2I(\sigma^2>0)$. 这时由定理 1.4.1，不论 z 的均值取什么值，z 总是正则的. 所以在经典统计中，难以见到“正则性”这个本质的假定.

第 2 章

线性统计问题

记号 2.0.1 在本章中，我们恒设 $z_T=(z_t,\ t\in T)$ 为一 m 维正则过程，它的生成空间为 $(\mathscr{H}(z_T),\ R)$，它的再生核表示空间为 $(\mathscr{R}(z_T),\ S_0)$，且令

$$\mathscr{H}^+(z_T)\triangleq C+\mathscr{H}(z_T)\triangleq\{c+x:\ c\in C,\ x\in\mathscr{H}(z_T)\}.$$

所谓基于观测过程 z_T 的线性统计问题，简单地说，就是限于取 $\mathscr{H}^+(z_T)$ 的元作为线性统计量，来尽可能"好"地预测与 z_T 有关的随机变量的值，或估计含于 z_T 均值中的未知参数，或作假设检验等等. 而所谓"好"的预测与估计，通常是用均方误差来衡量的.

§2.1 线性最小方差预报

定义 2.1.1 设 z_T 为一 m 维正则过程，$y\in\mathscr{L}_2^k$. 若存在 $\hat{y}\in\mathscr{H}^+(z_T)^k$ 使对任一 $x\in\mathscr{H}^+(z_T)^k$ 都有

$$R^+(y-\hat{y})\leqslant R^+(y-x)$$

成立，则称 $\hat{y}$ 为 y 基于 z_T 的线性最小方差预报，且记作

$$\hat{y}\triangleq\pi(y|z_T).$$

$R^+(y-\hat{y})\triangleq E(y-\hat{y})(y-\hat{y})^*$ 称为预报的均方误差(阵).

定理 2.1.2 令 $f_t\triangleq Ez_t^*,\ t\in T$. $\hat{y}$ 存在且唯一，它由下式表出：

$$\hat{y}=z(Z^y)-S_0(Z^y,\ f)+Ey,$$

$$R^+(y-\hat{y})=R(y-\hat{y})=R(y)-R(\hat{y})=R(y)-S_0(Z^y).$$

证 首先由系 1.2.10，$Z^y\in\mathscr{R}(z_T)^k$，它的逆再生表示像为 $z(Z^y)\in\mathscr{H}(z_T)^k$，即有 $R(z(Z^y),z_t)=R(y,\ z_t)$，$(t\in T)$成立.

由此易推出，对任意 $x\in\mathscr{H}^+(z_T)^k$ 有 $R(y-z(Z^v),\ x)=0$. 又由定理 1.2.12 知 $Ez(Z^v)=S_z(Z^v,\ f)$. 于是对任一 $x\in\mathscr{H}^+(z_T)^k$,

$$\begin{aligned}R^+(y-x)&=R(y-x)+(Ey-Ex)(Ey-Ex)^*\\&\geqslant R(y-z(Z^v)+z(Z^v)-x)\\&=R(y-z(Z^v))+R(z(Z^v)-x)\\&\geqslant R(y-z(Z^v))=R(y-z(Z^v),\ y)\\&=R(y)-S_z(Z^v).\end{aligned}$$

显然，当且仅当 $Ex=Ey$ 且 $R(z(Z^v)-x)=0$ 时，上式中两不等号取等号. 也就是说，使上式的均方误差达到最小的 $x\in\mathscr{H}^+(z_T)^k$，唯一地只能是 $x=z(Z^v)+c$(其中 $c\in C^k$)且 $Ex=S_z(Z^v,\ f)+c=Ey$. 由此即得所需结论. ∎

定理 2.1.3

i) $\hat{y}=\pi(y|z_T)$ 由下列三条件唯一确定：a) $\hat{y}\in\mathscr{H}^+(z_T)^k$; b) $E\hat{y}=Ey$; c) $R(y-\hat{y},\ z_t)=0$.

ii) 对任意 $y_1\in\mathscr{L}_2^k$, $y_2\in\mathscr{L}_2^l$, $C_1\in C^{n\times k}$, $C_2\in C^{n\times l}$ 有

$$\pi(C_1y_1+C_2y_2)=C_1\pi(y_1|z_T)+C_2\pi(y_2|z_T).$$

iii) 当 $y_n\xrightarrow{\mathscr{L}_2}y$(依 $\mathscr{L}_2$ 范数收敛)时，

$$\pi(y_n|z_T)\xrightarrow{\mathscr{L}_2}\pi(y|z_T).$$

证 i) 由定理 2.1.2 是显然的. ii) 由 i) 立即推出. 至于 iii)，由 ii)和定理 2.1.2 有

$$R(\hat{y}-\hat{y}_n)=R(\widehat{y-y_n})\leqslant R(y-y_n),$$

且 $E\hat{y}_n=Ey_n$, $E\hat{y}=Ey$. 由此即得所证. ∎

例 2.1.4 考虑 T 为单元集的情形. 设 z 为正则，此时由定理 1.4.1 和 2.1.2 有

$$\pi(y|z)=R(y,\ z)R(z)^-(z-Ez)+Ey,$$

$$R(y-\pi(y|z))=R(y)-R(y,\ z)R(z)^-R(z,\ y).$$

熟悉 Hilbert 空间初步的读者即可以看出，y 关于 z_T 的线性最小方差预报，当它们的均值都假设为零时(这时 z_T 当然是正则过程)，就是 y 在 Hilbert 空间 $(\mathscr{H}(z_T),\ R)$ 上的正交投影(严

格地说，$\hat{y}$ 的每个分量是 y 的相应分量在 $\mathscr{H}(z_T)$上的投影）．我们之所以在本节不厌其烦地重新加以讨论，其理由是：（一）我们并不以投影的存在性定理作为前提，而是用系 1.2.10 保证它的存在性；（二）将它推广到 z_T 的均值非 0 的情形(这时必须假定 z_T 为正则过程)；(三)更重要的是：定理 2.1.2 给出了 $\hat{y}$ 和 $R(\hat{y})$的再生核表示形式，以便于今后的实际计算(详见第三、四、六、七诸章)．

§ 2.2 回归系数的线性估计

设$(\Omega, \mathscr{F})$为一可测空间，$\{P_\theta: \theta\in C^p\}$ 是其上的一族概率测度，z_T 关于任一 P_θ 都是m维二阶过程，且进一步设

$$E_\theta z_t^* = \theta^* F_t,\ R_\theta(z_t, z_s) \equiv R(z_t, z_s),\ t,\ s\in T$$

而 $F\in \mathscr{R}(z_T)^p$．换句话说，z_T 满足如下的线性回归模型：

$$z_t = F_t^* \theta + v_t,\ t\in T. \tag{2.2.1}$$

其中 F 的 p 个行向量函数(均为 $\mathscr{R}(z_T)$中元)称为*回归变量*，θ 为 p 维未知参数(称为*回归系数*)向量，v_T 是一均值为 0、协方差已知且与 θ 无关的m维二阶过程．关于这一模型的进一步讨论，还可见后面的注记 2.2.7 和 2.2.10．

由 $F\in \mathscr{R}(z_T)^p$ 这一假设知 $\theta^* F\in \mathscr{R}(z_T)$，所以据定理 1.2.11，$z_T$ 关于任一 P_θ 都是m维正则过程，而协方差 R 与参数 θ 无关．于是由1.2.4可以定义 z_T 的统一的生成空间$(\mathscr{H}(z_T), R)$．

定义 2.2.1 对线性回归模型(2.2.1)，若存在 $x\in \mathscr{H}^+(z_T)^p$，使对任意的 $\theta\in C^p$ 恒有 $E_\theta x = \theta$ 成立，则称 x 为 θ 基于 z_T 的*线性无偏估计*．若 $\hat{\theta}$ 为 θ 的一个线性无偏估计，且对 θ 的任一线性无偏估计 x，恒有 $R(\hat{\theta})\leqslant R(x)$，则称 $\hat{\theta}$ 为 θ *基于 z_T 的线性最小方差无偏估计*，或称 Gauss–Markov 估计，简记作 GM 估计．

定理 2.2.2 GM 估计 $\hat{\theta}$ 存在的必要充分条件是，诸回归变量线性无关，而这一条件又等价于 $S_z(F)$ 为满秩阵．此时唯一地有

$$\hat{\theta} = S_z(F)^{-1} z(F),\ R(\hat{\theta}) = S_z(F)^{-1}.$$

证 首先证明定理所述的两个条件是等价的．若存在非零向量 $\theta_0\in C^p$ 使 $\theta_0^*F_t=0$, $t\in T$, 则 $\theta_0^*S_z(F)\theta_0=S_z(\theta_0^*F)=0$, 所以 $S_z(F)$是降秩的．反之，若 $S_z(F)$是降秩的，则存在 $0\neq\theta_0\in C^p$ 使 $0=\theta_0^*S_z(F)\theta_0=S_z(\theta_0^*F)$, 于是对任意 $t\in T$ 有

$$\|\theta_0^*F_t\|^2=\|R(z(\theta_0^*F),\ z_t)\|^2 \leqslant S_z(\theta_0^*F)\operatorname{tr}R(z_t)=0,$$

即 $\theta_0^*F=0$.

其次证明 $S_z(F)$满秩对于 $\hat\theta$ 的存在是充分必要的．

必要性：设存在 θ 的线性无偏估计 $x=c+x_0$, 其中

$$c\in C^p,\ x_0\in\mathscr{H}(z_T)^p.$$

则由定理 1.2.12,

$$\theta=E_\theta x=c+R(x_0,\ z(\theta^*F))=c+R(x_0,\ z(F))\theta$$

对任意 $\theta\in C^p$ 成立．于是推得 $c=0$, $R(x_0,\ z(F))=I$, 从而有 $x=x_0\in\mathscr{H}(z_T)^p$. 另一方面，由例 2.1.4 及引理 1.1.2 知，对任意 $x\in\mathscr{H}(z_T)^k$, $\theta\in C^p$,

$$\begin{aligned}\pi_\theta(x\,|\,z(F))&=R(x,\ z(F))S_z(F)^-(z(F)-E_\theta z(F))+E_\theta x\\&=R(x,\ z(F))S_z(F)^-(z(F)\\&\quad-S_z(F)\theta)+R(x,\ z(F))\theta\\&=R(x,\ z(F))S_z(F)^-z(F).\end{aligned}$$

特别要指出的是，此式表明 $\pi_\theta(x\,|\,z(F))$实际上与参数 θ 无关（今后记作 $\pi(x\,|\,z(F))$），因而可以当作一个新的统计量．由定理 2.1.2, 它的均值（对任一 P_θ）与 x 的相同，而方差（阵）却比 x 的要小．这一重要事实在本章中还将多次用到．

由上得知，如果存在 θ 的线性无偏估计 x, 则必有

$$\begin{aligned}I&=R(x,\ z(F))=R(\pi(x\,|\,z(F)),\ z(F))\\&=R(S_z(F)^-z(F),\ z(F))=S_z(F)^-S_z(F),\end{aligned}$$

所以 $S_z(F)$必为满秩阵，而 $S_z(F)^-=S_z(F)^{-1}$.

充分性：设 $S_z(F)$为满秩的．则

$$E_\theta S_z(F)^{-1}z(F)=S_z(F)^{-1}S_z(F)\theta=\theta,$$

且由必要性的证明知，对 θ 的任一线性无偏估计 x 有

$$R(x)\geqslant R(\pi(x|z(F))=R(S_z(F)^{-1}z(F))$$
$$=S_z(F)^{-1}S_z(F)S_z(F)^{-1}=S_z(F)^{-1},$$

所以 $\hat{\theta}=S_z(F)^{-1}z(F)$ 是 θ 的 GM 估计，且当且仅当 $x=\hat{\theta}$ 时，上式的不等号取等号(由§1投影的唯一性). ∎

有时，未知的回归系数向量 θ 并不取遍整个 C^p，而受某种已知的线性约束，这时 θ 估计的方差还可进一步缩小.

定义 2.2.3 设已给定 $L\in C^{l\times p}$, $b\in C^l$, 且设线性方程 $L\theta=b$ 在 C^p 中至少有一个解. 对线性回归模型(2.2.1)，如果存在 $x\in\mathscr{H}^+(z_T)^p$，使对满足 $L\theta=b$ 的一切 θ 恒有 $E_\theta x=\theta$ 成立，则称 x 为 θ 的带线性约束的线性无偏估计，对其中方差阵的最小者记作 $\hat{\hat{\theta}}$，称为 θ 的带线性约束的 GM 估计.

定理 2.2.4 设诸回归变量线性无关，则可表

$$\hat{\hat{\theta}}=\hat{\theta}-S_z(F)^{-1}L^*(LS_z(F)^{-1}L^*)^-(L\hat{\theta}-b),$$
$$R(\hat{\hat{\theta}})=S_z(F)^{-1}-S_z(F)^{-1}L^*(LS_z(F)^{-1}L^*)^-LS_z(F)^{-1},$$

其中 $\hat{\theta}=S_z(F)^{-1}z(F)$ 是 θ 的 GM 估计.

证 设 $x=c+x_0$ 是 θ 的一个带线性约束的无偏估计，其中 $c\in C^p$, $x_0\in\mathscr{H}(z_T)^p$. 类似于定理 2.2.2 证明中的讨论，如果 x_0 用 $\pi(x_0|z(F))=R(x_0,z(F))S_z(F)^{-1}z(F)=R(x_0,z(F))\hat{\theta}$ 来替代，那末 $x_1\triangleq c+\pi(x_0|z(F))$ 满足 $E_\theta x_1=E_\theta x$，且 $R(x_1)\leqslant R(x)$. 即 x_1 仍是 θ 的带约束的线性无偏估计，且其方差阵变小. 于是，为了求 $\hat{\hat{\theta}}$，不妨假定它具有形式 $\hat{\hat{\theta}}=c+(I_p-A)\hat{\theta}$（其中 $A\in C^{p\times p}$），然后定出 c 和 A. 为此，两边取期望得：对于 $L\theta=b$ 的任意解，有 $\theta=c+\theta-A\theta$ 成立，即 $A\theta=c$. 如令 θ_0 是 $L\theta=b$ 的某个确定的解，则有 $A(\theta-\theta_0)=0$. 换句话说，使 $L(\theta-\theta_0)=0$ 的任意 $\theta\in C^p$，都有 $A(\theta-\theta_0)=0$，即 L 的零化子空间含于 A 的零化子空间；而这又等价于 A 的每一行向量属于 L 的行向量张成的线性子空间，即存在 $B\in C^{p\times l}$ 使 $A=BL$. 因此 $c=A\theta_0=BL\theta_0=Bb$，而

$$\hat{\hat{\theta}}=Bb+(I-BL)\hat{\theta}=\hat{\theta}-B(L\hat{\theta}-b).$$

现在，只须由极小化 $R(\hat{\hat{\theta}})$ 来定出 B. 为简便起见，暂记

$S \triangleq S_e(F)$. 注意由

$$[L-(LS^{-1}L^*)(LS^{-1}L^*)^-L]S^{-1}[L-(LS^{-1}L^*)(LS^{-1}L^*)^-L]^*$$
$$=LS^{-1}L^*-2LS^{-1}L^*+LS^{-1}L^*=0,$$

知 $L=(LS^{-1}L^*)(LS^{-1}L^*)^-L$. 于是用配平方法, 可表

$$R(\hat{\hat{\theta}})=R((I-BL)\hat{\theta})=(I-BL)S^{-1}(I-BL)^*$$
$$=[B-S^{-1}L^*(LS^{-1}L^*)^-]LS^{-1}L^*$$
$$\times[B-S^{-1}L^*(LS^{-1}L^*)^-]^*$$
$$+S^{-1}-S^{-1}L^*(LS^{-1}L^*)^-LS^{-1}.$$

上式当 $B=S^{-1}L^*(LS^{-1}L^*)^-$ 时取极小, 而极小方差阵为 $R(\hat{\hat{\theta}})=S^{-1}-S^{-1}L^*(LS^{-1}L^*)^-LS^{-1}$. 值得注意的是: $\hat{\hat{\theta}}$ 本身也满足线性约束条件: $L\hat{\hat{\theta}}=b$ 这可由 $\hat{\hat{\theta}}$ 的表达式左乘 L 并利用 $(LS^{-1}L^*)\cdot(LS^{-1}L^*)^-L=L$ 而得到. ▌

定义 2.2.5 对线性回归模型(2.2.1), 设 E_{ap} 表示未知参数向量 θ 的某一先验概率分布的期望算子, 并且已知其一、二阶先验矩为

$$\mu_{ap}\triangleq E_{ap}\theta,\ \Gamma_{ap}\triangleq R_{ap}(\theta)=E_{ap}(\theta-\mu_{ap})(\theta-\mu_{ap})^*,$$

若存在 $\check{\theta}\in\mathscr{H}^+(z_T)^p$, 使对任意 $x\in\mathscr{H}^+(z_T)^p$ 恒有

$$E_{ap}E_\theta(\theta-\check{\theta})(\theta-\check{\theta})^*\leqslant E_{ap}E_\theta(\theta-x)(\theta-x)^*,$$

则称 $\check{\theta}$ 为 θ 的线性 Bayes 估计, 简称 LB 估计. 它的一个特点是不要求回归变量线性无关.

定理 2.2.6 $\check{\theta}$ 总是存在的, 一般可表为

$$\check{\theta}=\Gamma_{ap}S_s(F)(S_e(F)+S_s(F)\Gamma_{ap}S_s(F))^-(z(F)-S_s(F)\mu_{ap})$$
$$+\mu_{ap},$$
$$E_{ap}E_\theta(\theta-\check{\theta})(\theta-\check{\theta})^*$$
$$=\Gamma_{ap}-\Gamma_{ap}S_s(F)(S_e(F)+S_s(F)\Gamma_{ap}S_s(F))^-S_s(F)\Gamma_{ap},$$

若 Γ_{ap} 为满秩的, 则上两式可进一步简化成

$$\check{\theta}=(\Gamma_{ap}^{-1}+S_s(F))^{-1}(z(F)+\Gamma_{ap}^{-1}\mu_{ap}),$$
$$E_{ap}E_\theta(\theta-\check{\theta})(\theta-\check{\theta})^*=(\Gamma_{ap}^{-1}+S_s(F))^{-1}.$$

特别若取 $\Gamma_{ap}=\gamma I(\gamma>0)$, 则 $\check{\theta}$ 常称为 θ 的岭估计(通常取 $\mu_{ap}=0$), 这时若诸回归变量线性无关, 则当 $\gamma\to\infty$ 时,

$$\check{\theta} \xrightarrow{\text{a.s.}, \mathscr{L}_2} \hat{\theta}.$$

证 设 $x=c+x_0\in\mathscr{H}^+(z_T)^p$，其中 $c\in C^p$，$x_0\in\mathscr{H}(z_T)^p$，则由 $E_\theta x=c+R(x_0, z(F))\theta$ 知

$$\begin{aligned}
&E_{ap}E_\theta(\theta-x)(\theta-x)^* \\
&\quad=E_{ap}(\theta-c-R(x_0,z(F))\theta)(\theta-c-R(x_0,z(F))\theta)^*+R(x_0) \\
&\quad=(\mu_{ap}-c-R(x_0, z(F))\mu_{ap})(\mu_{ap}-c-R(x_0, z(F))\mu_{ap})^* \\
&\qquad+R_{ap}(\theta-R(x_0, z(F))\theta)+R(x_0).
\end{aligned}$$

同前，若用 $\pi(x_0|z(F))$ 代替 x_0，则上式前两项都不变而第三项 $R(\pi(x_0|z(F))\leqslant R(x_0)$，故 $\check{\theta}$ 应取 $c+Az(F)$ 的形式. 下面只须定出 c 和 $A\in C^{p\times p}$.

此时容易算出(暂记 $S\triangleq S_z(F)$)：

$$\begin{aligned}
&E_{ap}E_\theta(\theta-\check{\theta})(\theta-\check{\theta}) \\
&\quad=[(I-AS)\mu_{ap}-c][(I-AS)\mu_{ap}-c]^* \\
&\qquad+(I-AS)\Gamma_{ap}(I-AS)^*+ASA^* \\
&\quad=[(I-AS)\mu_{ap}-c][(I-AS)\mu_{ap}-c]^* \\
&\qquad+[A-\Gamma_{ap}S(S+S\Gamma_{ap}S)^-](S+S\Gamma_{ap}S) \\
&\qquad\times[A-\Gamma_{ap}S(S+S\Gamma_{ap}S)^-]^* \\
&\qquad+[\Gamma_{ap}-\Gamma_{ap}S(S+S\Gamma_{ap}S)^-S\Gamma_{ap}],
\end{aligned}$$

其中用到了 $S(S+S\Gamma_{ap}S)^-(S+S\Gamma_{ap}S)=S$ 这一事实. 这是因为任一使 $(S+S\Gamma_{ap}S)\theta=0$ 的 θ 必有 $\theta^*S\theta=0$，从而 $S\theta=0$，故可表 $S=R(S+S\Gamma_{ap}S)$.

上式当 $A=\Gamma_{ap}S(S+S\Gamma_{ap}S)^-$ 且 $c=\mu_{ap}-AS\mu_{ap}$ 时达到极小，且极小值即为其最后一项. 这就证得定理的一般公式.

若进一步设 Γ_{ap} 为满秩阵，则只须注意可表 $S+S\Gamma_{ap}S=S\Gamma_{ap}(\Gamma_{ap}^{-1}+S)=(\Gamma_{ap}^{-1}+S)\Gamma_{ap}S$，因此有

$$\begin{aligned}
&(I-AS)\Gamma_{ap}(I-AS)^*+ASA^* \\
&\quad=[A-(\Gamma_{ap}^{-1}+S)^{-1}](S+S\Gamma_{ap}S)[A-(\Gamma_{ap}^{-1}+S)^{-1}]^* \\
&\qquad+\Gamma_{ap}-\Gamma_{ap}S(\Gamma_{ap}^{-1}+S)^{-1},
\end{aligned}$$

故极小解为 $A=(\Gamma_{ap}^{-1}+S)^{-1}$，$c=\mu_{ap}-(\Gamma_{ap}^{-1}+S)^{-1}S\mu_{ap}=$

$(\Gamma_{ap}^{-1}+S)^{-1}(\Gamma_{ap}^{-1}+S-S)\mu_{ap}=A\Gamma_{ap}^{-1}\mu_{ap}$，而极小阵为

$$\Gamma_{ap}-\Gamma_{ap}S(\Gamma_{ap}^{-1}+S)^{-1}=(\Gamma_{ap}^{-1}+S)^{-1}.$$

定理的最后一个结论由上面的公式及定理2.2.2立即推出. ▮

注记2.2.7 本节前面部分的讨论都是针对模型(2.2.1)的. 当 z_T 的维数 $m=1$ 时，这一模型是常见且合理的；但当 $m>1$ 时，模型初看起来似乎不太合用. 因为 z_T 的各分量有不同的回归变量，即 F 的各列，但却有相同的回归系数 θ. 关于这一模型的直接应用，可见第四章§2.

另一类受到广泛重视的模型是多元统计分析中常见的[21]：

$$z_t=C^*\Theta g_t+v_t,\ t\in T, \tag{2.2.2}$$

其中 $C\in C^{q\times m}$ 是已知的矩阵，g 是 T 上的 C^p 值函数(回归变量)，$\Theta\in C^{q\times p}$ 是未知参数(回归系数)阵，而 v_T 是一均值为0、协方差已知且与 Θ 无关的 m 维二阶过程. 问题则是要依据对 z_T 的观测求 Θ 的“好的”线性估计.

下面我们要指出，模型(2.2.2)实际上是模型(2.2.1)的特殊情形，或者说，可以化为(2.2.1)的形式. 为此，我们须简单引述一下矩阵的 Kronecker 乘积的定义与性质. 为了一般化起见，下面恒讨论复值的情形，这可能在记号上要增加一点麻烦. 只对实值情形感兴趣的读者，只须在下面的讨论中把记号“*”理解为矩阵的转置，并把复共轭记号通通去掉就可以了.

设 $A=(a_{ij})\in C^{k\times q}$，$B=(b_{ij})\in C^{l\times p}$，则 $A\otimes B\in C^{kl\times qp}$ 由如下的分块阵形式定义：

$$A\otimes B\triangleq(a_{ij}\bar{B})_{1\le i\le k,1\le j\le q},$$

其中 $\bar{B}$ 表 B 的复共轭(注重与实矩阵情形的定义略有不同). 容易验证：

$$(A\otimes B)^*=A^*\otimes B^*,\ (A\otimes B)(C\otimes D)=AC\otimes BD$$

(当矩阵乘法 AC、BD 可行时).

又对 $\Theta=(\theta_{ij})\in C^{q\times p}$，记

$$\Theta_v\triangleq(\theta_{11},\ \cdots,\ \theta_{1p},\ \theta_{21},\ \cdots,\ \theta_{q1},\ \cdots,\ \theta_{qp})^\tau\in C^{qp}$$

(τ 表示转置),称为 Θ 的行化向量,且容易验证:

$$(A\Theta B^{\tau})_r=(A\otimes B)\Theta_r.$$

根据上面这些关系式,对模型(2.2.2),可表

$$E_\Theta z_t=C^*\Theta g_t=(C^*\Theta g_t)_r=(C\otimes g_t)^*\Theta_r,\ (t\in T).$$

这样就化成了模型(2.2.1),而对 z_T 的正则性假设,现在变成 $C\otimes g\in\mathscr{R}(z_T)^{qp}$. 然后利用定理 2.2.2、2.2.4 和 2.2.6 的结果,立即可以给出(2.2.2)中未知参数阵 Θ 的(即 Θ_r 的)各种线性估计,例如:

定理 2.2.8 在正则性假设 $C\otimes g\in\mathscr{R}(z_T)^{pq}$ 下,Θ 的 GM 估计 $\hat{\Theta}$ 存在的必要充分条件是 $S_z(C\otimes g)$ 为满秩阵. 此时唯一地有

$$\hat{\Theta}_r=S_z(C\otimes g)^{-1}z(C\otimes g),\ R(\hat{\Theta}_r)=S_z(C\otimes g)^{-1}.$$

又 $S_z(C\otimes g)$ 为满秩的一个充分条件是 g 的 p 个分量线性无关,且 $\operatorname{rank}C=q(\leqslant m)$.

又如,考虑 Θ 有线性约束 $L\Theta M^*=B$,其中 L、M、B 为给定的矩阵,且方程至少有一个解. 由前述性质,这一约束可表成 $(L\otimes M)\Theta_r=B_r$. 于是有

定理 2.2.9 设 $S_z(C\otimes g)$ 为满秩阵,则 Θ 的带线性约束 $L\Theta M^*=B$ 的 GM 估计 $\hat{\hat{\Theta}}$ 可表为:

$$\hat{\hat{\Theta}}_r=\hat{\Theta}_r-S_z(C\otimes g)^{-1}(L\otimes M)^*[(L\otimes M)S_z(C\otimes g)^{-1}(L\otimes M)^*]^-\times[(L\otimes M)\hat{\Theta}_r-B_r],$$

$$R(\hat{\hat{\Theta}}_r)=S_z(C\otimes g)^{-1}-S_z(C\otimes g)^{-1}(L\otimes M)^*\times[(L\otimes M)S_z(C\otimes g)^{-1}(L\otimes M)^*]^-(L\otimes M)S_z(C\otimes g)^{-1}.$$

同样还可给出 Θ 的线性 Bayes 估计 $\check{\Theta}$.

注记 2.2.10 关于模型(2.2.2),应用较普遍的一种情形是 z_T 的协方差函数满足 $R(z_s,z_t)=R(v_s,v_t)=\lambda(s,t)\Gamma$, $(s,t\in T)$,其中 Γ 为 m 阶半正定阵,λ 为 T^2 上的复(或实)值函数. 这时,上述结果还可进一步简化,并可直接表出 $\hat{\Theta}$ 和 $\hat{\hat{\Theta}}$. 为此,记 Γ 的全体非零特征值与相应的单位正交特征向量分别为 $\gamma_1,\cdots,\gamma_l$; $v_1,\cdots,v_l$,并记

$$V \triangleq (\gamma_1^{-\frac{1}{2}}v_1, \cdots, \gamma_l^{-\frac{1}{2}}v_l) \in C^{m\times l},$$

则易知有 $V^*\Gamma V = I_l$（I_l 表 l 阶单位阵）$\Gamma VV^*\Gamma = \Gamma$ 成立，再令 $y_t \triangleq V^* z_t$, $t\in T$, 则

$$R(\bar{y}_s, \bar{y}_t) = \lambda(s, t) I_l,$$

$$R(y_s, z_t) = \bar{\lambda}(s, t) V^*\Gamma = R(y_s, \Gamma V y_t).$$

用 $(\mathscr{R}_\lambda, S_\lambda)$ 表示以 λ 为核的再生核空间，并设 $g\in\mathscr{R}_\lambda^q$，且 g 的诸分量线性无关，$C = \tilde{C}\Gamma$，且 $\mathrm{rank}\, C = q$. 又用 $\bar{y}^{(i)}(g)$（$i=1, \cdots, l$）表示 g 在 $\bar{y}_T$ 的第 i 分量 $\bar{y}_T^{(i)}$ 的生成空间 $\mathscr{H}(\bar{y}_T^{(i)})^p$ 中的逆再生表示，并记 $p\times l$ 随机矩阵 $\bar{Y}(g) \triangleq (\bar{y}^{(1)}(g), \cdots, \bar{y}^{(l)}(g))$. 显然有关系式：

$$R(y^{(i)}(g), y_t^{(j)}) = \delta_{ij}\bar{g}_t,$$

$$R(y^{(i)}(g), y^{(j)}(g)) = \delta_{ij}\overline{S_\lambda(g)}, \quad (t\in T;\ i, j=1, \cdots, l).$$

于是对任意 $t\in T$ 有

$$\begin{aligned}
&R((CV\bar{Y}(g)^*)_r, z_t) \\
&= R((CV\otimes I_p)(\bar{Y}(g)^*)_r, \Gamma V y_t) \\
&= (CV\otimes I_p)\begin{pmatrix} R(y^{(1)}(g), y_t) \\ \vdots \\ R(y^{(l)}(g), y_t) \end{pmatrix} V^*\Gamma \\
&= (CV\otimes I_p)(I_l\otimes g_t)(V^*\Gamma\otimes I_1) \\
&= (\tilde{C}\Gamma VV^*\Gamma)\otimes g_t = (\tilde{C}\Gamma)\otimes g_t = C\otimes g_t,
\end{aligned}$$

即得 $C\otimes g\in\mathscr{R}(z_T)^{qp}$，且 $z(C\otimes g) = (CV\bar{Y}(g)^*)_r$,

$$\begin{aligned}
S_z(C\otimes g) &= R(z(C\otimes g)) = R((CV\otimes I_p)(\bar{Y}(g)^*)_r) \\
&= (CV\otimes I_p)(I_l\otimes S_\lambda(g))(V^*C^*\otimes I_p) \\
&= (C\Gamma^- C^*)\otimes S_\lambda(g).
\end{aligned}$$

再利用定理 2.2.8、2.2.9 及 $(A\Theta B^*)_r = (A\otimes B)\Theta_r$，立刻推出

$$\begin{aligned}
\hat{\Theta} &= (C\Gamma^- C^*)^{-1} CV\bar{Y}(g)^* S_\lambda(g)^{-1}, \\
\hat{\hat{\Theta}} &= \hat{\Theta} - (C\Gamma^- C^*)^{-1} L^*[L(C\Gamma^- C^*)^{-1} L^*]^- \\
&\quad \times (L\hat{\Theta}M^* - B)[MS_\lambda(g)^{-1}M^*]^- MS_\lambda(g)^{-1}, \\
R(\hat{\Theta}_r) &= (C\Gamma^- C^*)^{-1}\otimes S_\lambda(g)^{-1},
\end{aligned}$$

$$R(\hat{\hat{\theta}}_r)=(C\Gamma^-C^*)^{-1}\otimes S_\lambda(g)^{-1}$$
$$-(C\Gamma^-C^*)^{-1}L^*[L(C\Gamma^-C^*)^{-1}L^*]^-L(C\Gamma^-C^*)^{-1}$$
$$\otimes S_\lambda(g)^{-1}M^*[MS_\lambda(g)^{-1}M^*]^-MS_\lambda(g)^{-1}.$$

注记2.2.11 在讨论模型(2.2.1)和(2.2.2)时，我们都假定 z_T 的协方差 R 是已知的. 但对许多实际问题，R 却并非已知或含未知参数，习惯上称为*多余参数*. 通常有两种情况：

i) $R=\sigma^2R^\triangle$，其中 $R^\triangle$ 为已知，而 $\sigma^2(>0)$ 为未知参数. 这时 R 与 $R^\triangle$ 是相互控制的，从而由定理1.3.3知，$\mathscr{R}(z_T)=\mathscr{R}^\triangle(z_T)$，且由定义可知，对任一 $F\in\mathscr{R}(z_T)^p=\mathscr{R}^\triangle(z_T)^p$ 有

$$S_z^\triangle(F)=\sigma^2S_z(F),\quad z^\triangle(F)=\sigma^2z(F).$$

于是，对模型(2.2.1)，定理2.2.2关于 GM 估计的结果可改写为 $\hat{\theta}=S_z^\triangle(F)^{-1}z^\triangle(F)$（与 σ^2 无关），$R(\hat{\theta})=\sigma^2S_z^\triangle(F)^{-1}$. 定理2.2.4的结果也可作相应的改写. 类似地，对模型(2.2.2)，定理2.2.8、2.2.9也有相应的改变.

ii) R 未知，但假设它被一已知的协方差 $R^\triangle$ 控制. 这时，由定理1.3.3，$\mathscr{R}(z_T)\subset\mathscr{R}^\triangle(z_T)$. 对模型(2.2.1)和(2.2.2)，分别设 $F\in\mathscr{R}^\triangle(z_T)^p$ 和 $C\otimes g\in\mathscr{R}^\triangle(z_T)^{qp}$. 若 F 和 $C\otimes g$ 的行向量函数分别是线性无关的，则可给出 θ 和 Θ 的一种估计：

$$\theta^\triangle\triangleq S_z^\triangle(F)^{-1}z^\triangle(F)\ (\in\mathscr{H}^\triangle(z_T)^p),$$
$$\Theta_r^\triangle\triangleq S_z^\triangle(C\otimes g)^{-1}z^\triangle(C\otimes g)\ (\in\mathscr{H}^\triangle(z_T)^{qp}),$$

它们仍然是线性无偏估计，但一般不再是最小方差的. 若

$$R(x)\leqslant\beta R^\triangle(x),\quad x\in\mathscr{H}_0(z_T),$$

则可给出估计的方差阵及其上界：

$$R(\theta^\triangle)=S_z^\triangle(F)^{-1}R(z^\triangle(F))S_z^\triangle(F)^{-1}\leqslant\beta S_z^\triangle(F)^{-1},$$
$$R(\Theta_r^\triangle)=S_z^\triangle(C\otimes g)^{-1}R(z^\triangle(C\otimes g))S_z^\triangle(C\otimes g)^{-1}$$
$$\leqslant\beta S_z^\triangle(C\otimes g)^{-1}.$$

例2.2.12 考虑 T 为单元集的情形. 这时模型(2.2.1)成为

$$z=F^*\theta+v.$$

由定理1.4.1，关于 z 的正则性假设就是 F 的每一行向量都属于由 $R(z)$ 的行向量张成的线性空间. 若 $R(z)$ 为满秩，则这一假

设对任一 $F\in C^{p\times m}$ 都成立. 在正则性假设下,

$$S_z(F)=FR(z)^-F^*,\ z(F)=FR(z)^-z.$$

于是, 若 $\operatorname{rank} F=p$(即 F 的 p 个行向量线性无关), 则 θ 的 GM 估计为

$$\hat{\theta}=(FR(z)^-F^*)^{-1}FR(z)^-z,\quad R(\hat{\theta})=(FR(z)^-F^*)^{-1}.$$

这是经典统计学中熟知的结果. 若 $R(z)=\sigma^2 I_m$, 则 GM 估计就是最小二乘估计. 又因为总存在 $\alpha>0$ 使 $R(z)\leqslant\alpha I_m$, 所以即使取消正则性假设, 只要 $\operatorname{rank} F=p$, θ 的最小二乘估计总是存在的, 而且是线性无偏的(相当于取 $R^{\Delta}(z)=I_m$).

现在, 设 $T=\{1,\ \cdots,\ n\}$, 考虑模型(2.2.2):

$$z_t=C^*\Theta g_t+v_t,\ (t=1,\ \cdots,\ n),$$

记 $$Z\triangleq(z_1,\ \cdots,\ z_n)\in C^{m\times n},\ V\triangleq(v_1,\ \cdots,\ v_n)\in C^{m\times n},$$

$$G\triangleq(g_1,\ \cdots,\ g_n)\in C^{p\times n}$$

则上式可表成矩阵形式:

$$Z=C^*\Theta G+V.$$

进一步利用注记 2.2 7 开头引述的定义和性质, 即有

$$Z_r=(C\otimes G)^*\Theta_r+V_r.$$

这同前一例子的模型是一样的. 于是, 若 $C\otimes G$ 的行向量都属于 $R(Z_r)$ 的行向量张成的线性空间, 且 $\operatorname{rank}(C\otimes G)=qp$, 则 Θ 的 GM 估计为

$$\hat{\Theta}_r=[(C\otimes G)R(Z_r)^-(C\otimes G)^*]^{-1}(C\otimes G)R(Z_r)^-Z_r,$$

$$R(\hat{\Theta}_r)=[(C\otimes G)R(Z_r)^-(C\otimes G)^*]^{-1}.$$

若进一步设 $R(z_s,\ z_t)=\sigma^2\bar{\lambda}_{st}\Gamma$, $(s,\ t\in T)$, $\Lambda\triangleq(\lambda_{st})$, σ^2 可以是未知的. 则易知有 $R(Z_r)=\sigma^2\Gamma\otimes\Lambda$, $R(Z_r)^-=\sigma^{-2}\Gamma^-\otimes\Lambda^-$. 此时, 如果 C 和 G 的行向量分别属于 Γ 和 Λ 的行向量张成的线性空间, 且 $\operatorname{rank} C=q$, $\operatorname{rank} G=p$, 则有

$$\begin{aligned}R(\hat{\Theta}_r)&=\sigma^2[(C\otimes G)(\Gamma^-\otimes\Lambda^-)(C\otimes G)^*]^{-1}\\&=\sigma^2(C\Gamma^-C^*)^{-1}\otimes(G\Lambda^-G^*)^{-1},\end{aligned}$$

$$\begin{aligned}\hat{\Theta}_r&=[(C\Gamma^-C^*)^{-1}\otimes(G\Lambda^-G^*)^{-1}](C\Gamma^-\otimes G\Lambda^-)Z_r\\&=[(C\Gamma^-C^*)^{-1}C\Gamma^-\otimes(G\Lambda^-G^*)^{-1}G\Lambda^-]Z_r.\end{aligned}$$

后一式再表回矩阵形式，即得

$$\hat{\Theta}=(C\Gamma^{-}C^{*})^{-1}C\Gamma^{-}Z\Lambda^{-}G^{*}(G\Lambda^{-}G^{*})^{-1}.$$

这是经典统计学中一个较新的结果.

同样，在最后所述的条件下，容易算出Θ的带线性约束$L\Theta M^{*}=B$的 GM 估计为：

$$\begin{aligned}\hat{\hat{\Theta}}=&\hat{\Theta}-(C\Gamma^{-}C^{*})^{-1}L^{*}(L(C\Gamma^{-}C^{*})^{-1}L^{*})^{-}(L\hat{\Theta}M^{*}-B)\\&\times(M(G\Lambda^{-}G^{*})^{-1}M^{*})^{-}M(G\Lambda^{-}G^{*})^{-1},\\R(\hat{\hat{\Theta}}_{r})=&\sigma^{2}(C\Gamma^{-}C^{*})^{-1}\otimes(G\Lambda^{-}G^{*})^{-1}\\&-\sigma^{2}(C\Gamma^{-}C^{*})^{-1}L^{*}(L(C\Gamma^{-}C^{*})^{-1}L^{*})^{-}L(C\Gamma^{-}C^{*})^{-1}\\&\otimes(G\Lambda^{-}G^{*})^{-1}M^{*}(M(G\Lambda G^{*})^{-1}M^{*})^{-}M(G\Lambda^{-}G^{*})^{-1}\end{aligned}$$

§2.3 可估性及其推广

在本节中仍讨论模型(2.2.1). 从定理2.2.2可知道，在$F\in\mathscr{R}(z_T)^{p}$这一正则性假设下，若F的p个行向量函数(回归变量)线性无关(等价于$S_c(F)$为p阶满秩阵)，则θ的 GM 估计$\hat{\theta}$存在. 这时，若$a(\theta)\triangleq A\theta$, 其中$A\in C^{k\times p}$, 则$\hat{a}\triangleq A\hat{\theta}$显然是$a$关于$z_T$的线性无偏估计，而且是最小方差的(见系2.3.3). 但是在一些有重要实际意义的场合，回归变量是线性相关的. 这时，$\hat{\theta}$将不再存在，但对有些$a(\theta)$, $\hat{a}$却仍可能存在. 这样的$a(\theta)$在经典统计学中称为(k维)可估函数. 现在我们将这一概念稍加推广，并给出关于可估性的必要充分条件.

定义2.3.1 设z_T满足线性回归模型(2.2.1)，而

$$y=A\theta+u(\in\mathscr{L}_{1}^{k}),\tag{2.3.1}$$

其中$A\in C^{k\times p}$为已知矩阵，$u\in\mathscr{L}_{20}^{k}$, 且u的(即y的)方差阵及其与z_T的协方差阵都是已知的(与θ无关). 若存在$x\in\mathscr{H}^{+}(z_T)^{k}$, 使对任意的$\theta\in C^{p}$恒有$E_\theta x=E_\theta y=A\theta$成立，则称$y$为可估的，称$x$为$y$关于$z_T$的线性无偏估计. 若$\hat{y}$是$y$的一个线性无偏估计，且对$y$的任意线性无偏估计$x$都有$R(y-\hat{y})\leqslant R(y-x)$，则

称 $\hat{y}$ 为 y 的 Gauss-Markov 估计(GM 估计).

如果未知参数向量 θ 还满足相容的线性约束条件 $L\theta=b$, 其中 $L\in C^{l\times p}$, $b\in C^l$, 则类似地可定义 y 的可估性为：存在 $x\in \mathscr{H}^+(z_T)^k$, 使对一切满足 $L\theta=b$ 的 θ 恒有 $E_\theta x=A\theta$ 成立，且相应地可定义 y 的带线性约束的 GM 估计 $\hat{\hat{y}}$

注意：这里的 $\hat{y}$ 和 $\hat{\hat{y}}$ 与 2.1.1 中定义的线性最小方差预报不同，因为 z_T 与 y 的均值中都含有未知参数 θ（在经典统计中，$\hat{y}$ 和 $\hat{\hat{y}}$ 仍称为预报，这里为区别起见，改称为估计）. 又若(2.3.1)中的 $u=0$, 则定义和下述定理就化为 θ 的可估函数及其估计(这时 $Z^y=0$).

定理 2.3.2　在相容线性约束 $L\theta=b$ 下，y 可估的必要充分条件是：存在 $C\in C^{k\times p}$、$D\in C^{k\times l}$ 使 $A=CS_z(F)+DL$. 此时，y 的带线性约束的 GM 估计为

$$\hat{\hat{y}}=z(Z^y)+(A-S_z(Z^y,\ F))[P_L(P_LS_z(F)P_L)^-P_L \times(z(F)-S_z(F)L^*(LL^*)^-b)+L^*(LL^*)^-b],$$

$$R(y-\hat{\hat{y}})=R(y)-S_z(Z^y)+(A-S_z(Z^y,\ F)) \times P_L(P_LS_z(F)P_L)^-P_L(A-S_z(Z^y,\ F))^*,$$

其中
$$P_L\triangleq I_p-L^*(LL^*)^-L.$$

证　在证明中暂记 $S\triangleq S_z(F)$.

必要性：设 $x=c+x_0$ 为 y 的一个带约束的线性无偏估计，其中 $c\in C^k$, $x_0\in\mathscr{H}(z_T)^k$. 则对满足 $L\theta=b$ 的一切 θ, 恒有

$$A\theta=E_\theta y=E_\theta x=c+R(x,\ z(F))\theta.$$

取 θ_0 为 $L\theta=b$ 的一个确定的解，则上式表明，对任意满足 $L(\theta-\theta_0)=0$ 的 $\theta\in C^p$, 恒有 $(A-R(x,\ z(F)))(\theta-\theta_0)=0$, 而这又等价于存在 $D\in C^{k\times l}$, 使 $A-R(x,\ z(F))=DL$. 由引理 1.1.2,

$$R(x,\ z(F))=R(x,\ z(F))S^-S,$$

故有　$A=R(x,\ z(F))+DL=R(x,\ z(F))S^-S+DL.$

充分性：设 $A=CS+DL$. 由于显然有 $P_L^*=P_L$, $P_L^2=P_L$(即 P_L 为投影阵), $(LP_L)(LP_L)^*=LP_LL^*=0$, 因而 $LP_L=0$;

$$[P_L-P_L(P_LSP_L)^-P_LSP_L]^*S[P_L-P_L(P_LSP_L)^-P_LSP_L]=0,$$

因而 $SP_L=SP_L(P_LSP_L)^-P_LSP_L$（从投影的角度看，这些等式都是明显的）．由此知 $AP_L=AP_L(P_LSP_L)^-P_LSP_L$，$S_s(Z^y, F)P_L=R(y, z(F))S^-SP_L=S_s(Z^y, F)P_L(P_LSP_L)^-P_LSP_L$．于是，若 $\hat{\hat{y}}$ 如定理所设，则对任意满足 $L\theta=b$ 的 $\theta\in C^p$，有：

$$\begin{aligned}E_\theta\hat{\hat{y}}&=S_s(Z^y, F)\theta+(A-S_s(Z^y, F))\\&\quad\times[P_L(P_LSP_L)^-P_LSP_L\theta+L^*(LL^*)^-L\theta]\\&=S_s(Z^y, F)\theta+(A-S_s(Z^y,F))\cdot\\&\quad\times(P_L+L^*(LL^*)^-L)\theta=A\theta.\end{aligned}$$

即 $\hat{\hat{y}}$ 是 y 的一个带约束的无偏估计，故 y 是可估的．此时，若 $x\in\mathscr{H}^+(z_T)^k$ 是 y 的任一带约束的线性无偏估计，注意必要性部分的证明及 $R(y-z(Z^y), z(Z^y))=0=R(y-z(Z^y), x)$，就有

$$\begin{aligned}R(y-x)&=R(y-z(Z^y))+R(x-z(Z^y))\\&\geqslant R(y-z(Z^y))+R(\pi(x-z(Z^y)\mid z(F)))\\&=R(y)-S_s(Z^y)+R(x-y, z(F))\\&\quad\times S^-SS^-R(z(F), x-y)\\&\geqslant R(y)-S_s(Z^y)+R(x-y, z(F))\\&\quad\times S^-SP_L(P_LSP_L)^-P_LSS^-R(z(F), x-y)\\&=R(y)-S_s(Z^y)+(A-S_s(Z^y, F))\\&\quad\times P_L(P_LSP_L)^-P_L(A-S_s(Z^y, F))^*\\&=R(y-z(Z^y))+R(\hat{\hat{y}}-z(Z^y))=R(y-\hat{\hat{y}}).\end{aligned}$$

所以 $\hat{\hat{y}}$ 确是 y 的带约束的 GM 估计．其中第二个不等号用到了不等式 $S\geqslant SP_L(P_LSP_L)^-P_LS$，这可由

$$\begin{aligned}0&\leqslant(I-SP_L(P_LSP_L)^-P_L)S(I-SP_L(P_LSP_L)^-P_L)^*\\&=S-SP_L(P_LSP_L)^-P_LS\end{aligned}$$

得出（也可由 $S^{1/2}P_L(P_LSP_L)^-P_LS^{1/2}$ 为投影阵因而 $\leqslant I$ 明显看出）．于是定理得证．∎

系 2.3.3 i）无约束条件时，y 可估的必要充分条件是：存在 $C\in C^{k\times p}$，使 $A=CS_s(F)$．此时 y 的 GM 估计为

$$\hat{y}=z(Z^y)+(A-S_s(Z^y, F))S_s(F)^-z(F),$$

$$R(y-\hat{y})=R(y)-S_s(Z^y)$$
$$+(A-S_s(Z^y,F))S_s(F)^-(A-S_s(Z^y,F))^*.$$

ii) 若回归变量线性无关,则任何形如(2.3.1)的 y 都是可估的,且对无约束和带约束的情形分别有

$$\hat{y}=z(Z^y)+(A-S_s(Z^y,F))\hat{\theta},$$
$$\hat{\hat{y}}=z(Z^y)+(A-S_s(Z^y,F))\hat{\hat{\theta}}.$$

其中 $\hat{\theta}$、$\hat{\hat{\theta}}$ 分别由定理 2.2.2、2.2.4 给出. 特别此时有恒等式

$$R(\hat{\hat{\theta}})=P_L(P_LS_s(F)P_L)^-P_L$$
$$=S_s(F)^{-1}-S_s(F)^{-1}L^*(LS_s(F)^{-1}L^*)^-LS_s(F)^{-1}.$$

证 i) 无约束条件即相当于定理 2.3.2 中的 $L=0$, $b=0$. 此时 $P_L=I$.

ii) 由定理 2.2.2 知, 回归变量线性无关等价于 $S_s(F)$ 为满秩阵. 这时定理 2.3.2 中的条件自然满足(可取 $C=AS_s(F)^{-1}$, $D=0$). 至于所列的矩阵恒等式,可以从对比定理 2.2.4 和 2.3.2 的结果得出. 下面给出一个简单而有趣的纯矩阵证明:

设 S、T 为两个 p 阶满秩对称阵,令

$$U\triangleq S^{-1}-S^{-1}L^*(LS^{-1}L^*)^-LS^{-1},$$
$$V\triangleq T^{-1}-T^{-1}L^*(LT^{-1}L^*)^-LT^{-1}.$$

($T=I_p$ 时, $V=P_L$), 则极易算出

$$USV=VSU=V,\ UTV=VTU=U.$$

于是得

$$V(VSV)^-V=USV(VSV)^-VSU$$
$$=UTVSV(VSV)^-VSVTU$$
$$=UTVSVTU=USVTU=VTU=U. \blacksquare$$

定义 2.3.4 设 y 满足(2.3.1),其中的未知参数 θ 有已知的先验一、二阶矩: $E_{op}\theta=\mu_{op}$, $R_{op}(\theta)=\Gamma_{op}$. 若存在 $\check{y}\in\mathscr{H}^+(z_T)^k$, 使对任意 $x\in\mathscr{H}^+(z_T)^k$ 恒有

$$E_{op}E_\theta(y-\check{y})(y-\check{y})^*\leqslant E_{op}E_\theta(y-x)(y-x)^*,$$

则称 $\check{y}$ 为 y 的线性 Bayes 估计,简称 LB 估计.

定理 2.3.5 y 的 LB 估计恒存在,且可表为

$$\check{y}=z(Z^v)+(A-S_z(Z^v,\ F))\check{\theta},$$

$$E_{\alpha p}E_\theta(y-\check{y})(y-\check{y})^*$$
$$=R(y)-S_z(Z^v)+(A-S_z(Z^v,\ F))$$
$$\times[E_{\alpha p}E_\theta(\theta-\check{\theta})(\theta-\check{\theta})^*](A-S_z(Z^v,\ F))^*,$$

其中 $\check{\theta}$ 为 θ 的 LB 估计，其表示式见定理 2.2.6. 特别，若 $\Gamma_{\alpha p}=\gamma I$，$y$ 为可估的，则当 $\gamma\to\infty$ 时，$\check{y}\xrightarrow{a.s.,\mathscr{L}_2}\hat{y}$.

证　仿定理 2.2.6 和 2.3.2 即可证得. ▌

另外，按照注记 2.2.7 的讨论，我们也可就模型(2.2.2)讨论随机矩阵 $Y=A\Theta B^*+U$(其中 U 是均值为 0 的随机阵)的可估性问题，并容易给出 Y 可估的必要充分条件及相应的 GM 估计与 LB 估计. 在此不再赘述.

§2.4　最大信噪比问题

定义 2.4.1　设 m 维二阶过程 z_T 可表为如下形式：

$$z_t=f_t^*+g_t^*+v_t,\ (t\in T),\tag{2.4.1}$$

其中 v_T 均值为 0，而 f、$g\in\mathscr{H}(z_T)$，且 $f\neq0$. 在实际应用中，常把 f^* 理解为 z_T 中含有的确定信号，g^*+v_T 理解为 z_T 中含有的噪声. 由假设知 z_T 是一正则过程.

设 $0\neq x\in\mathscr{H}(z_T)$，由定理 1.2.12,

$$Ex=R(x,\ z(f+g))=R(x,\ z(f))+R(x,\ z(g)).$$

在电学中常用 $|R(x,\ z(f))|^2$ 和 $E|x-R(x,\ z(f))|^2=|R(x,\ z(g)|^2+R(x)$ 分别表示 x 含有的平均信号能量和噪声能量(因为能量与电流或电压的平方成正比). 记

$$r_x\triangleq\frac{|R(x,\ z(f))|^2}{|R(x,\ z(g))|^2+R(x)},$$

称为 x 的信噪比. 若存在 $x_0\in\mathscr{H}(z_T)$，使

$$r_{x_0}=\sup_{x\in\mathscr{H}(z_T)}r_x,$$

则称 x_0 为 z_T 的最大信噪比线性统计量，r_{x_0} 为最大信噪比.

定理 2.4.2 一般有

$$x_0=\lambda\left(z(f)-\frac{S_z(f,\ g)}{1+S_z(g)}\ z(g)\right),$$

$$r_{x_0}=S_z(f)-\frac{|S_z(f,\ g)|^2}{1+S_z(g)},$$

其中 λ 为任一非零常数. 特别是若 $S_z(f,\ g)=0$ (更特别, 若 $g=0$), 则

$$x_0=\lambda z(f),\ r_{x_0}=S_z(f).$$

证 首先, 由 r_x 的定义显然有 $r_{\lambda x}=r_x$. 其次, 若记

$$\hat{x}\triangleq\pi(x\,|\,z(f),\ z(g)),$$

则由于

$$R(\hat{x},\ z(f))=R(x,\ z(f)),$$

$$R(\hat{x},\ z(g))=R(x,\ z(g)),$$

而 $R(\hat{x})\leqslant R(x)$, 所以 $r_{\hat{x}}\geqslant r_x$, 并且当且仅当 $\hat{x}=x$ 或 $R(x,\ z(f))=0$ 时有 $r_{\hat{x}}=r_x$ 成立, 而后一情形显然不可能使 r_x 为最大 (注意 $f\neq 0$). 于是, 在考虑最大信噪比问题时, 不妨限于取 x 为 $z(f)$ 与 $z(g)$ 的线性组合. 更进一步, 若令 $h\triangleq g-\alpha^* f$, 其中

$$\alpha=S_z(f,\ g)/S_z(f),$$

则有 $S_z(h,\ f)=R(z(h),\ z(f))=0,\quad z(g)=\alpha^* z(f)+z(h),$

$$S_z(g)=R(z(g))=|\alpha|^2 S_z(f)+S_z(h).$$

基于上述理由, 不妨限于考虑 $x=z(f)+\beta z(h)$, 其中 β 为待定系数. 这时,

$$r_x=\frac{|S_z(f)|^2}{|\alpha S_z(f)+\beta S_z(h)|^2+S_z(f)+|\beta|^2 S_z(h)},$$

要 r_x 为最大, 必须且只需上式分母为最小. 将其配平方, 得

$$\text{分母}=S_z(h)(1+S_z(h))\left|\ \beta+\frac{\alpha S_z(f)}{1+S_z(h)}\right|^2$$

$$+S_z(f)+\frac{|\alpha|^2 S_z(f)^2}{1+S_z(h)}.$$

可见, 当且仅当 $h=0$ 或 $\beta=-\dfrac{\alpha S_z(f)}{1+S_z(h)}$ 时分母为最小. 于是

$$
\begin{aligned}
x_0 &= z(f)-\frac{\alpha S_z(f)}{1+S_z(h)}z(h)\\
&= z(f)-\frac{\alpha S_z(f)}{1+S_z(g)-|\alpha|^2S_z(f)}(z(g)-\alpha^* z(f))\\
&= \frac{1+S_z(g)}{1+S_z(g)-|\alpha|^2S_z(f)}\left[z(f)-\frac{S_z(f,g)}{1+S_z(g)}z(g)\right],\\
\sigma_{x_0} &= S_z(f)\Big/\left[1+\frac{|\alpha|^2S_z(f)}{1+S_z(h)}\right]\\
&= S_z(f)(1+S_z(h))/(1+S_z(h)+|\alpha|^2S_z(f))\\
&= S_z(f)(1+S_z(g)-|\alpha|^2S_z(f))/(1+S_z(g))\\
&= S_z(f)-\frac{|S_z(f,g)|^2}{1+S_z(g)}.
\end{aligned}
$$

即所证. ▌

关于本问题的一种特殊情形的讨论，可见谢衷洁，程乾生的文章[17],[18]. 还可见汪嘉冈的本章[19].

§2.5 正态过程分布间的绝对连续性条件

本章前四节讨论的内容都是有关随机过程的统计估计问题. 在本节中，我们将讨论与随机过程的统计假设检验问题有关的某些结果. 假设检验的目的是要判断一个观测到的样本是否服从某个确定的概率分布. 尤其是对于简单假设检验，即判断样本服从两个确定概率分布中的哪一个时，常用到的基本统计量是这两个分布的似然比或对数似然比. 对于有限样本而言，它们就是两个相应的有限维概率密度之比或其对数. 但当样本大小为无穷时，因为无穷维 Lebesque 测度并不存在，通常意义的概率密度也就不存在. 这时就需要直接考虑两个概率测度之间的绝对连续性问题，而把它们的 Radon-Nikodym 导数(参见定义 2.5.1)定义为似然比.

例如，我们观测的对象是随机过程 $z_T=(z_t,\ t\in T)$，现在要根据观测来判断其中是否包含某一确定的信号 $f_t=Ez_t,\ t\in T$；或

者 $Ez_t=0$. 这在通讯理论中称为信号检测问题. 为了进行似然比检验, 首先就要考虑两个相应的概率分布在什么条件下绝对连续, 然后求出其 Radon-Nikodym 导数.

随机过程概率分布之间的绝对连续性的研究内容十分丰富. 近年来, 随着现代鞅论的进展已经获得很多深入的结果. 在本节中, 只是对多维正态过程的情形作一些讨论, 而且不打算涉及对假设检验问题的具体应用, 其主要目的是为了从另一方面揭示再生表示和正则性的意义. 读者跳过本节并不影响以后的阅读.

定义 2.5.1 设 $P^\triangle$ 和 P 是可测空间 $(\Omega, \mathscr{F})$ 上的两个概率测度. 若任意使 $P^\triangle(A)=0$ 的 $A\in\mathscr{F}$, 恒有 $P(A)=0$, 则称 P 对 $P^\triangle$ **绝对连续**, 记作 $P\ll P^\triangle$. 若 $P\ll P^\triangle$ 和 $P^\triangle\ll P$ 同时成立, 则称 P 与 $P^\triangle$ **相互绝对连续**, 记作 $P\equiv P^\triangle$. 若存在 $A\in\mathscr{F}$ 使 $P^\triangle(A)=0$ 且 $P(A)=1$, 则称 P 与 $P^\triangle$ **正交或相互奇异**, 记作 $P\perp P^\triangle$.

由著名的 Radon-Nikodym 定理, $P\ll P^\triangle$ 的必要充分条件是: 存在对 $P^\triangle$ 可积的非负随机变量 $\frac{dP}{dP^\triangle}$, 使对一切 $A\in\mathscr{F}$,

$$P(A)=\int_A \frac{dP}{dP^\triangle}\,dP^\triangle \triangleq E^\triangle\left(\frac{dP}{dP^\triangle}1_A\right), \qquad (2.5.1)$$

其中 1_A 是 A 的示性函数. $\frac{dP}{dP^\triangle}$ 称为 P 对 $P^\triangle$ 的 Radon-Nikodym 导数, 它在 a.s. $P^\triangle$ 的意义下是唯一的. 如果还有 $\frac{dP}{dP^\triangle}>0$, a.s. P 成立, 则 $P\equiv P^\triangle$, 且 $\frac{dP^\triangle}{dP}=1\Big/\frac{dP}{dP^\triangle}$, a.s. P, $P^\triangle$. 又若存在序列 $(A_n)\subset\mathscr{F}$, 使 $P^\triangle(A_n)\to0$, $P(A_n)\to1$, 则 $P\perp P^\triangle$. 因为这时可取子序列 (A_{n_k}) 使 $P^\triangle(A_{n_k})<\frac{1}{2^k}$, 再令 $A\triangleq\bigcap_l\bigcup_{k\geq l}A_{n_k}$, 则显然有 $P^\triangle(A)=0$, $P(A)=1$.

下面讨论 m 维实正态过程 z_T 在样本空间上导出的概率分布之间绝对连续和正交的条件. 为此, 我们在本节对 C 理解为实数全体, 而样本空间就是 $((C^m)^T, \mathscr{B}((C^m)^T))$, 其中 $(C^m)^T$ 表示 T

上的 C^m 值函数全体，$\mathscr{B}((C^m)^T)$ 表示 $(C^m)^T$ 上的全体有限维柱集生成的 σ 域. 所谓正态过程，就是所有有限维分布都服从正态分布的过程.

引理 2.5.2 设 P^Δ 和 P 分别为 m 维实正态过程 z_T 在样本空间上导出的两个概率分布，其中 $E^\Delta z_t=0$, $Ez_t=f_t^*$, $t\in T$. 其协方差分别用 R^Δ 和 R 表示，相应的再生核空间分别用 $(\mathscr{R}^\Delta(z_T), S_z^\Delta)$ 和 $(\mathscr{R}(z_T), S_z)$ 表示. 若 R 与 R^Δ 不相互控制，或虽然 R 与 R^Δ 相互控制，但 $f\bar{\in}\mathscr{R}(z_T)$，则 $P\perp P^\Delta$.

证 若 R^Δ 不控制 R. 由定义 1.3.1 易知，必存在 $(x_n)\subset \mathscr{H}_0(z_T)$，使 $R^\Delta(x_n)=1$，而 $\sigma_n^2\triangleq R(x_n)\to\infty$. 令

$$A_n\triangleq\{|x_n|>\sqrt{\sigma_n}\}.$$

因 $E^\Delta x_n=0$，故由 Chebyshev 不等式有

$$P^\Delta(A_n)\leqslant\frac{R^\Delta(x_n)}{\sigma_n}=\frac{1}{\sigma_n}\to 0.$$

又注意 x_n 是正态随机变量，故

$$P(A_n)=1-\frac{1}{\sqrt{2\pi}\sigma_n}\int_{-\sqrt{\sigma_n}}^{\sqrt{\sigma_n}}\exp\left\{-\frac{(u-Ex_n)^2}{2\sigma_n^2}\right\}du$$

$$\geqslant 1-\frac{1}{\sqrt{2\pi}}\int_{-1/\sqrt{\sigma_n}}^{1/\sqrt{\sigma_n}}e^{-\frac{v^2}{2}}dv\to 1.$$

因此 $P\perp P^\Delta$. 同样，若 R 不控制 R^Δ，则存在 $(y_n)\subset\mathscr{H}_0(z_T)$，使 $R(y_n)=1$ 而 $\delta_n^2\triangleq R^\Delta(x_n)\to\infty$. 令

$$B_n\triangleq\{|y_n-Ey_n|\leqslant\sqrt{\delta_n}\},$$

同理可证 $P^\Delta(B_n)\to 0$, $P(B_n)\to 1$. 因此也有 $P\perp P^\Delta$.

现在设 R 与 R^Δ 相互控制，即存在 $0<\beta_1<\beta_2<\infty$，使对任意 $x\in\mathscr{H}_0(z_T)$ 有 $\beta_1R(x)\leqslant R^\Delta(x)\leqslant\beta_2R(x)$. 由定理 1.3.3 知 $\mathscr{R}(z_T)=\mathscr{R}^\Delta(z_T)$. 若 $f\bar{\in}\mathscr{R}(z_T)$，则由定理 1.2.9，存在

$$\left(x_n\triangleq\sum_i c_i^{(n)*}z_{t_i^{(n)}}\right)\subset\mathscr{H}_0(z_T),$$

使 $R(x_n)=1$，而 $0<Rx_n=\left(\sum_i f_{t_i^{(n)}}c_i^{(n)}\right)^*\to\infty$. 令

$$A_n\triangleq\left\{x_n>\frac{1}{2}Ex_n\right\},$$

由 Chebyshev 不等式，即得

$$P^{\Delta}(A_n)\leqslant P^{\Delta}\left\{|x_n|>\frac{1}{2}Ex_n\right\}\leqslant\frac{4R^{\Delta}(x_n)}{(Ex_n)^2}\leqslant\frac{4\beta_2}{(Ex_n)^2}\to 0,$$

$$\begin{aligned}P(A_n)&=P\left\{x_n-Ex_n>-\frac{1}{2}Ex_n\right\}\\&\geqslant P\left\{|x_n-Ex_n|\leqslant\frac{1}{2}Ex_n\right\}\\&\geqslant 1-\frac{4}{(Ex_n)^2}\to 1.\ \blacksquare\end{aligned}$$

定理 2.5.3　假设同引理 2.5.2，且进一步设 $R=R^{\Delta}$，则 $P\equiv P^{\Delta}$ 的必要充分条件为 $f\in\mathscr{R}(z_T)$。此时有

$$\frac{dP}{dP^{\Delta}}=\exp\left\{z(f)-\frac{1}{2}S_z(f)\right\}.$$

否则，$P\perp P^{\Delta}$.

证　必要性与"否则"部分由引理 2.5.2 即得.

充分性：若 $f\in\mathscr{R}(z_T)$，则定理中给出的 $\frac{dP}{dP^{\Delta}}$ 自然是一个非负随机变量. 我们只需用它来验证(2.5.1)式. 显然，又只须对所有有限维柱集 A 来进行验证. 再由有限维分布与其特征函数的一一对应关系，则只须证明，对任意有限集 $\{t_j\}\subset\boldsymbol{T}$, $\{c_j\}\subset C^n$ 有

$$E^{\Delta}\left(\frac{dP}{dP^{\Delta}}\exp\{i\sum c_j^*z_{t_j}\}\right)=E\exp\{i\sum c_j^*z_{t_j}\}$$

成立.

将定理给出的 $\frac{dP}{dP^{\Delta}}$ 代入上式左边，并注意 $\sum c_j^*z_{t_j}$, $z(f)$ 都是正态随机变量，由正态变量指数函数的期望公式即得

$$\begin{aligned}左边&=\exp\left\{-\frac{1}{2}S_z(f)\right\}E^{\Delta}\exp\{z(f)+i\sum c_j^*z_{t_j}\}\\&=\exp\left\{-\frac{1}{2}S_z(f)\right\}\exp\left\{\frac{1}{2}E^{\Delta}(z(f)+i\sum c_j^*z_{t_j})^2\right\}\\&=\exp\left\{-\frac{1}{2}S_z(f)\right\}\exp\left\{\frac{1}{2}[S_z(f)-R(\sum c_j^*z_{t_j})+2i\sum f_{t_j}c_j]\right\}\end{aligned}$$

$$=\exp\left\{iE(\sum c_j^* z_{t_j})-\frac{1}{2}R(\sum c_j^* z_{t_j})\right\}=E\exp\{i\sum c_j^* z_{t_j}\}.$$

于是 $P\ll P^{\Delta}$. 又显然有 $\frac{dP}{dP^{\Delta}}>0$, a. s. P^{Δ}, P, 所以 $P^{\Delta}\ll P$. 故 $P\equiv P^{\Delta}$. ▌

下一定理是定理 2.5.3 的推广.

定理 2.5.4 假设同引理 2.5.2. 则 $P\equiv P^{\Delta}$ 的必要充分条件为以下三条件同时成立:

i) R 与 R^{Δ} 相互控制;

ii) $f\in\mathscr{R}(z_T)$;

iii) 由 $R(x, y)=R^{\Delta}(x, By)$, x, $y\in\mathscr{H}(z_T)$ 确定的线性有界正算子 B, 使 $I-B$ 为($\mathscr{H}(z_T)$, R^{Δ})上的 Hilbert-Schmidt 算子(简称 HS 算子). 其中 I 表 $\mathscr{H}(z_T)$上的单位算子.

此时, 若以(λ_n)和(e_n)分别表 B 的非 1 本征值列和相应的标准正交本征元列,则有

$$\frac{dP}{dP^{\Delta}}=\exp\left\{z(f)-\frac{1}{2}S_{\epsilon}(f)-\frac{1}{2}\sum_{n\geq 1}[(\lambda_n^{-1}-1)e_n^2+\ln\lambda_n]\right\}.$$

如上述三条件有一个不成立,则 $P\perp P^{\Delta}$.

证 首先, 由引理 2.5.2, 条件 i) 和 ii)对 $P\equiv P^{\Delta}$ 是必要的; 而且只要它们中的一个不成立, 就有 $P\perp P^{\Delta}$. 因此, 我们不妨在假设 i)和 ii)成立的条件下来证明条件 iii)的充分必要性. 其次, 若令 P' 是 z_T 的另一正态概率分布: $E'z_T=0$, $R'=R$. 则由定理 2.5.3 知 $P\equiv P'$. 于是,如果已知 $P'\equiv P^{\Delta}$, 则有 $P\equiv P^{\Delta}$, 且

$$\frac{dP}{dP^{\Delta}}=\frac{dP}{dP'}\,\frac{dP'}{dP^{\Delta}};$$

如果已知 $P'\perp P^{\Delta}$, 则有 $P\perp P^{\Delta}$. 换句话说, 根据已得结果, 我们又不妨在 $f=0$(即 $Ez_T=0$)的假定下来证明定理.

充分性: 由 i), $\mathscr{H}(z_T)=\mathscr{H}^{\Delta}(z_T)$, 且存在 $0<\beta_1<\beta_2<\infty$, 使对任意 $x\in\mathscr{H}(z_T)$ 有 $\beta_1 R(x)\leqslant R^{\Delta}(x)\leqslant\beta_2 R(x)$. 又定理 1.3.2 保证了正算子 B 的存在性. 由 iii), $I-B$ 为 HS 算子. 由 HS 算子、本征值及标准正交本征元的定义, 且注意 z_T 对 P^{Δ} 和 P

都是均值为 0 的正态过程，即易得知以下诸性质：$\sum\limits_{n\geqslant1}(1-\lambda_n)^2<\infty$，$(e_n)$ 对 P^Δ 和 P 都是均值为 0 的、相互独立的正态随机变量列，$R^\Delta(e_n)=1$，$R(e_n)=R^\Delta(e_n,\ Be_n)=R^\Delta(e_n,\ \lambda_n e_n)=\lambda_n$，$\beta_1\leqslant\lambda_n^{-1}\leqslant\beta_2$，$n\geqslant1$。又对任意 $x\in\mathscr{H}(z_T)$，若记 $x_0\triangleq x-\sum\limits_{n\geqslant1}R^\Delta(x,\ e_n)e_n$，则有 $(I-B)x_0=0$，即 $Bx_0=x_0$，因而 $R(x_0)=R^\Delta(x_0)$，且 x_0 对 P^Δ、P 都独立于 (e_n)。

分别表

$$
\begin{aligned}
&(\lambda_n^{-1}-1)e_n^2+\ln\lambda_n\\
&\quad=(\lambda_n^{-1}-1)(e_n^2-1)+(1-\lambda_n+\ln\lambda_n)+\lambda_n^{-1}(1-\lambda_n)^2\\
&\quad=(\lambda_n^{-1}-1)(e_n^2-\lambda_n)+(1-\lambda_n+\ln\lambda_n).
\end{aligned}
$$

由以上所述诸性质知

$$
\begin{aligned}
&\sum\lambda_n^{-1}(1-\lambda_n)^2\leqslant\beta_2\sum(1-\lambda_n)^2<\infty,\\
&\sum|1-\lambda_n+\ln\lambda_n|\leqslant\beta_2^2\sum(1-\lambda_n)^2<\infty,\\
&\sum E^\Delta(\lambda_n^{-1}-1)^2(e_n^2-1)^2\\
&\quad=2\sum(\lambda_n^{-1}-1)^2\leqslant2\beta_2^2\sum(1-\lambda_n)^2<\infty,\\
&\sum E(\lambda_n^{-1}-1)^2(e_n^2-\lambda_n)^2\\
&\quad=2\sum\lambda_n^2(\lambda_n^{-1}-1)^2=2\sum(1-\lambda_n)^2<\infty.
\end{aligned}
$$

于是由关于独立和的 Kolmogorov 定理，随机级数 $\sum\limits_{n\geqslant1}[(\lambda_n^{-1}-1)e_n^2+\ln\lambda_n]$ a. s. P^Δ，P 收敛。

仿定理 2.5.3 的证法，对任意有限集 $\{t_j\}\subset T$，$\{c_j\}\subset C^m$，令 $x\triangleq\sum\limits_j c_j^*z_{t_j}$，则由 $x=\sum\limits_{n\geqslant1}R^\Delta(x,\ e_n)e_n+x_0$ 知

$$
\begin{aligned}
&E^\Delta\exp\left\{ix-\frac{1}{2}\sum_{n\geqslant1}[(\lambda_n^{-1}-1)e_n^2+\ln\lambda_n]\right\}\\
&\quad=\prod_{n\geqslant1}E^\Delta\exp\left\{iR^\Delta(x,\ e_n)e_n-\frac{1}{2}(\lambda_{n}^{-1}-1)e_n^2\right.\\
&\qquad\left.+\frac{1}{2}\ln\lambda_n\right\}E^\Delta e^{ix_0}\\
&\quad=\prod_{n\geqslant1}\frac{1}{\sqrt{2\pi\lambda_n}}\int_{-\infty}^{\infty}\exp\{iR^\Delta(x,\ e_n)u\}
\end{aligned}
$$

$$\times \exp\left\{-\frac{u^2}{2\lambda_n}\right\} du \cdot E^\Delta e^{ix_0}$$

$$=\prod_{n\geq 1} \exp\left\{-\frac{1}{2}\lambda_n R^\Delta(x, e_n)^2\right\} e^{-\frac{1}{2}R^\Delta(x_0)}$$

$$=\prod_{n\geq 1} E \exp\{iR^\Delta(x, e_n)e_n\} E e^{ix_0}$$

$$=E\exp\{i(\sum_{n\geq 1} R^\Delta(x, e_n)e_n + x_0)\} = Ee^{ix}.$$

于是 $P\ll P^\Delta$,

$$\frac{dP}{dP^\Delta}=\exp\left\{-\frac{1}{2}\sum_{n\geq 1}[(\lambda_n^{-1}-1)e_n^2+\ln\lambda_n]\right\}.$$

又因 $\frac{dP}{dP^\Delta}>0$, 所以 $P^\Delta\ll P$.

必要性: 设 $I-B$ 不是 HS 算子. 则由自伴算子的谱分解易知, 对任意 $n\geq 1$, 存在 $(\mathscr{H}(z_T), R^\Delta)$ 中的标准正交族 $\{e_{n1}, \cdots, e_{nK_n}\}$, 使 $Be_{nk}=\lambda_{nk}e_k$, 而 $\sum_{k=1}^{K_n}(1-\lambda_{nk})^2\geq n$. 这时 $\{e_{n1}, \cdots, e_{nK_n}\}$ 对 P^Δ 和 P 都是均值为 0 且相互独立的正态随机变量族,

$$R^\Delta(e_{nk})=1,\ R(e_{nk})=\lambda_{nk},\ \beta_1\leq\lambda_{nk}^{-1}\leq\beta_2.$$

令
$$A_n\triangleq\left\{\sum_{k=1}^{K_n}(\lambda_{nk}^{-1}-1)e_{nk}^2\leq\frac{1}{2}\sum_{k=1}^{K_n}(\lambda_{nk}^{-1}-\lambda_{nk})\right\},$$

则

$$P^\Delta(A_n)=P^\Delta\left\{\sum_k(\lambda_{nk}^{-1}-1)(e_{nk}^2-1)\leq-\frac{1}{2}\sum_k\lambda_{nk}^{-1}(1-\lambda_{nk})^2\right\}$$

$$\leq P^\Delta\left\{\left|\sum_k(\lambda_{nk}^{-1}-1)(e_{nk}^2-1)\right|\geq\frac{\beta_1}{2}\sum_k(1-\lambda_{nk})^2\right\}$$

$$\leq 8\left(\frac{\beta_2}{\beta_1}\right)^2\frac{1}{n}\to 0,$$

$$P(A_n)=P\left\{\sum_k(\lambda_{nk}^{-1}-1)(e_{nk}^2-\lambda_{nk})\leq\frac{1}{2}\sum_k\lambda_{nk}^{-1}(1-\lambda_{nk})^2\right\}$$

$$\geq P\left\{\left|\sum_k(\lambda_{nk}^{-1}-1)(e_{nk}^2-\lambda_{nk})\right|\leq\frac{\beta_1}{2}\sum_k(1-\lambda_{nk})^2\right\}$$

$$\geq 1-\frac{8}{n\beta_1^2}\to 1.$$

因此 $P\perp P^\Delta$. ▌

本节的定理是[13]中的结果对多维情形的推广.

在本章中，相当完备地讨论了任意指标集 $\boldsymbol{T}$ 上的多维二阶正则过程 z_T 的各类线性统计问题. 其中最值得注意的是：除了定理 2.5.4 之外，所有统计量都仅和 z_T 的再生核表示空间及其元的逆再生表示有关. 换句话说，随机过程 z_T 的各类线性统计问题的解可统一归结为：由 z_T 的协方差阵函数($\Gamma(s, t)\triangleq R(z_s, z_t)$, $s, t\in \boldsymbol{T}$) 定出 z_T 的再生核表示空间 $(\mathscr{R}(z_T), S_s)$，以及对任意 $F\in \mathscr{R}(z_T)^k$ 求出其逆再生表示 $z(F)\in \mathscr{H}(z_T)^k$. 把这些结果代入本章的诸定理，就得出了各种线性统计问题的具体答案. 而对一些常用的情形求出再生核表示空间及逆表示的算法，正是本书其余各章要研究的主要内容.

由第一章§4和本章诸例还可看出，对于 $\boldsymbol{T}$ 为单元集或有限集的情形，即经典统计的情形，线性统计问题的解本质上是求协方差阵 $R(z)$ 的弱逆(或逆阵) $R(z)^-$(或 $R(z)^{-1}$)的问题. 还须注意，当 $R(z)$ 非满秩时，我们只要求定出 $R(z)$ 的随便一个弱逆，而并不需要它在 Moore–Penrose 意义下的唯一的广义逆(即放弃 $R(z)^-R(z)$, $R(z)R(z)^-$ 为对称阵的条件). 这将显著减少计算量. 例如二阶半正定阵 $\begin{pmatrix}1 & 1\\ 1 & 1\end{pmatrix}$ 的 Moore–Penrose 广义逆为 $\begin{pmatrix}\frac{1}{4} & \frac{1}{4}\\ \frac{1}{4} & \frac{1}{4}\end{pmatrix}$，这当然也是它的弱逆；但它的一个更易求的弱逆却是 $\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}$. 一般，设 $H=\begin{pmatrix}H_{11} & H_{12}\\ H_{12}^* & H_{22}\end{pmatrix}$ 为一半正定 Hermite 阵，令

$$\widetilde{H}_{22}\triangleq H_{22}-H_{12}^*H_{11}^-H_{12},$$

$$H^-=\begin{pmatrix}H_{11}^-+H_{11}^-H_{12}\widetilde{H}_{22}^-H_{12}^*H_{11}^- & -H_{11}^-H_{12}\widetilde{H}_{22}^-\\ -\widetilde{H}_{22}^-H_{12}^*H_{11}^- & \widetilde{H}_{22}^-\end{pmatrix}, \quad (2.5.2)$$

则容易验证，这里的 H^- 必为 H 的一个弱逆(注意由引理 1.1.2 易知 $H_{12}=H_{11}H_{11}^-H_{12}$). 这样就把求 H^- 的问题化为求较低阶半

正定 Hermite 阵的弱逆问题，而最终归结为求一非负实数 h 的弱逆：$h^{-}=0$，若 $h=0$；$h^{-}=\frac{1}{h}$，若 $h>0$. 像 2.5.2 这样的简单等式，作者还未曾在别处见到过(通常都要求 H_{11}, $\tilde{H}_{22}$ 的逆阵存在)。

第 3 章

随机序列与协方差可分解情形

第一、二章的内容是对一般指标集 T 上的二阶过程 z_T 讨论的．为了定出 z_T 的再生核表示空间及其逆再生表示，从而求得各类线性统计问题的实际算法，就必须分别考虑各种类型的随机过程，首先是指标集 T 的具体化．最常见的指标集有两种：一种是 T 为整数集，相应的过程通常称为离散时间过程，特别当 T 为全体正整数 N 或全体整数时，又称为随机序列或时间序列；一种是 T 为实轴上某一有穷或无穷区间，相应的过程通常称为连续时间过程．其实，指标不一定代表时间，尤其有限指标集也可以是标志抽样或观测的序号．不过在很多常用的情形，T 的确是时间的标志．最后，还有一种以某一确定集合的子集为指标的随机过程（随机集函数）是在理论上较常用的，我们将在第五章加以讨论．

本章专门讨论 T 为有限集和 $T=N$ 的离散时间过程，这在应用上是最常见的，其数学处理也比较容易．

§3.1 一般递推解

记号 3.1.1 对每一 $t\in N$，记 $N_t\triangleq\{s\in N;\ s\leqslant t\}=\{1,\cdots,t\}$．设 $z_N=(z_s,\ s\in N)$ 为一 m 维二阶随机序列，它在 N_t 上的局限为 $z_{N_t}=\{z_1,\cdots,z_t\}$，当其为正则时，其生成空间 $\mathscr{H}(z_{N_t})$ 记作 $\mathscr{H}(z,t)$，或简记作 $\mathscr{H}_t$（显然，当 $s\leqslant t$ 时，$\mathscr{H}_s\subset\mathscr{H}_t$）；其再生核表示空间 $\mathscr{R}(z_{N_t})$ 记作 $\mathscr{R}(z,t)$，或简记作 $\mathscr{R}_t$．显然，若 $F\in\mathscr{R}_t^*$，则对任何 $s<t$，F 在 N_s 上的局限属子 $\mathscr{R}_s^*$，若 $F\in\mathscr{R}(z_N)^*$，则对任何 $t\in N$，F 在 N_t 上的局限属于 $\mathscr{R}_t^*$．今后如无混淆，N 上的

任一矩阵值函数 $F \triangleq (F_s,\ s\in N)$ 在 N_t 上的局限仍用 F 表示. 若 F 和 G 在 N_t 上的局限分别属于 $\mathscr{R}_t^k$ 和 $\mathscr{R}_t^l$, 则此二局限的内积阵记作 $S_z(F, G)_t$ 或简记作 $S(F, G)_t$, $S(F)_t \triangleq S(F, F)_t$, 而 F 在 $\mathscr{H}_t^k$ 中的逆再生表示则记作 $z(F)_t$.

注意, 若分别引进 mt 维随机向量, $k\times mt$ 及 $l\times mt$ 矩阵

$$Z=\begin{pmatrix} z_1 \\ \vdots \\ z_t \end{pmatrix},\ F=(F_1,\ \cdots,\ F_t),\ G=(G_1,\ \cdots,\ G_t),$$

则由定义及定理 1.4.1 易知 $\mathscr{H}_t=\mathscr{H}(Z)$; $F\in\mathscr{R}_t^k$ 当且仅当 $F\in\mathscr{R}(Z)^k$, 即当且仅当存在 $\tilde{F}\in C^{k\times mt}$ 使 $F=\tilde{F}R(Z)$. 此时

$$Z(F)_t=Z(F)=FR(Z)^-Z,$$

$$S(F,\ G)_t=S_z(F,\ G)=FR(Z)^-G^*.$$

换句话说, 可以把指标集为有限集的情形归结为单元集的情形, 只是把维数m放大了 t 倍. 但这样做在计算时须求一个 mt 阶半正定阵 $R(Z)$ 的弱逆; 随着 t 的增大, 计算量将迅猛增加. 尤其是当我们希望对每个 $t\in N$ 依次求出上述解答时, 这样做就更是不可取的. 合理的办法是设法给出 $z(F)_t$ 与 $S(F,\ G)_t$ 的递推方程, 这在电子计算机应用中尤其重要.

为下面叙述和书写的简便, 我们约定: 在本章中凡出现带指标 t 的随机变量或矩阵, 当 $t\leqslant 0$ 时, 都一律不言而喻地理解为 0 (标量、向量或矩阵).

定义 3.1.2 设 z_N 为一 m 维二阶随机序列. 若对任意 $t\in N$, z_{N_t} 都是正则的, 则称 z_N 为*局部正则序列*(自然, 均值为 0 的二阶序列总是局部正则的). 此时, 令

$$\varepsilon_t \triangleq z_t - z(Z^{z_t})_{t-1},\ (t\in N), \tag{3.1.1}$$

称 ε_N 为 z_N 的*新息序列*. 由定理 2.1.2, 除可能相差一常向量外, $z(Z^{z_t})_{t-1}$ 为 z_t 基于 $z_{N_{t-1}}$ 的线性最小方差预报, 因此 ε_t 可理解为从 z_t 中去掉其由从前的观察值可说明的部分后所余的"新的信息".

注意: 按前面的约定, $\varepsilon_1=z_1$. 又根据第一章和本章规定的

记号，当 t 固定时，Z^{z_t} 是 N 上的 m 阶矩阵值函数：

$$Z_s^{z_t} \triangleq R(z_t, z_s), \ (s \in N),$$

而 $z(Z^{z_t})_s$ 则是 Z^{z_t}(局限于 N_s)在 $\mathscr{H}_s^m$ 中的再生逆表示，即满足 $z(Z^{z_t})_s \in \mathscr{H}_s^m$，且对任意 $r \in N_s$ 有下式成立：

$$R(z(Z^{z_t})_s, z_r) = Z_r^{z_t} = R(z_t, z_r).$$

引理 3.1.3 设 z_N 为一 m 维局部正则序列.

i) 若对任意 $t \in N$ 有 $u_t \in \mathscr{H}_t^k$，则 u_N 是 k 维局部正则序列.

ii) 设 ε_N 是 z_N 的新息序列，则 $\varepsilon_t \in \mathscr{H}_t^m$，$\varepsilon_N$ 为 m 维局部正则序列，且对任意 $y \in \mathscr{H}_{t-1}^l$ 有 $R(\varepsilon_t, y) = 0$. 特别对 t、$s \in N$，$t \neq s$ 有 $R(\varepsilon_t, \varepsilon_s) = 0$ 及

$$R(\varepsilon_t) = R(\varepsilon_t, z_t) = R(z_t) - S(Z^{z_t})_{t-1}. \qquad (3.1.2)$$

证 i) 设 $f_t = Ez_t^*$，$(t \in N)$. 则由定理 1.2.12，

$$Eu_s^* = R(z(f)_t, u_s) \triangleq U_s^{z(f)_t}, \ (s \in N_t).$$

由系 1.2.10，$U^{z(f)_t} \in \mathscr{R}(u, t)$. 故由定理 1.2.11 知 u_N 是局部正则的.

ii) 结论的前半部分由 i)得出. 又由定义 3.1.2 的最后一式知 $R(\varepsilon_t, z_s) = 0$，$(s \in N_{t-1})$，由此易得结论的后半部分. ▌

注 ε_N 的性质 $R(\varepsilon_t, \varepsilon_s) = 0 (t \neq s)$ 常称为*正交性*，具此性质的随机序列称为*正交随机序列*，技术科学中又习称为*白噪声*(*序列*)，否则，称为*非白噪声*或*有色噪声*. 新息序列 ε_N 相当于 z_N 的 Schmidt 正交化，它在今后的讨论中起着根本性的作用.

定理 3.1.4 设 z_N 为一 m 维局部正则序列，$F \triangleq (F_s, s \in N)$ 为 N 上的 $C^{k \times m}$ 值函数. 则 F (在 N_t 上的局限) $\in \mathscr{R}_t^k$ 的必要充分条件是 $F \in \mathscr{R}_{t-1}^k$ 且存在 $C_t \in C^{k \times m}$ 使

$$F_t - S(F, Z^{z_t})_{t-1} = C_t R(\varepsilon_t).$$

此时，对任意 $F \in \mathscr{R}_t^k$，$G \in \mathscr{R}_t^l$，有下列递推表达式：

$$z(F)_t = z(F)_{t-1} + (F_t - S(F, Z^{z_t})_{t-1}) R(\varepsilon_t)^- \varepsilon_t, \qquad (3.1.3)$$

$$S(F, G)_t = S(F, G)_{t-1} + (F_t - S(F, Z^{z_t})_{t-1}) R(\varepsilon_t)^- \times (G_t - S(G, Z^{z_t})_{t-1})^*, \qquad (3.1.4)$$

$$\varepsilon_t = z_t - z(Z^{z_t})_{t-1},\quad R(\varepsilon_t) = R(z_t) - S(Z^{z_t})_{t-1}.$$

此外还有

$$F_t - S(F,\ Z^{z_t})_{t-1} = R(z(F)_t,\ \varepsilon_t). \qquad (3.1.5)$$

成立.

证 必要性：设 $F\in \mathscr{R}_t^k$, 则存在唯一的 $z(F)_t\in \mathscr{H}_t^k$ 使

$$F_s = R(z(F)_t,\ z_s),\ s\in N_t.$$

若考虑 F 在 N_{t-1} 上的局限，则由系 1.2.10 知 $F\in\mathscr{R}_{t-1}^k$. 于是又存在唯一的 $z(F)_{t-1}\in\mathscr{H}_{t-1}^k$, 使

$$F_s = R(z(F)_{t-1},\ z_s),\ (s\in N_{t-1}).$$

由上两式易推得，对任意 $y\in\mathscr{H}_{t-1}^1$ 有

$$R(z(F)_t,\ y) = R(z(F)_{t-1},\ y)$$

成立. 因此由引理 1.1.2 有

$$\begin{aligned}
&F_t - S(F,\ Z^{z_t})_{t-1}\\
&= R(z(F)_t,\ z_t) - R(z(F)_{t-1},\ z(Z^{z_t})_{t-1})\\
&= R(z(F)_t,\ z_t) - R(z(F)_t,\ z(Z^{z_t})_{t-1}) = R(z(F)_t,\ \varepsilon_t)\\
&= R(z(F)_t,\ \varepsilon_t)R(\varepsilon_t)^- R(\varepsilon_t) \triangleq C_t R(\varepsilon_t).
\end{aligned}$$

充分性：设 $F\in\mathscr{R}_{t-1}^k$ 且存在 $C_t\in \boldsymbol{C}^{k\times m}$ 使

$$F_t - S(F,\ Z^{z_t})_{t-1} = C_t R(\varepsilon_t).$$

令 $\qquad x \triangleq z(F)_{t-1} + (F_t - S(F,\ Z^{z_t})_{t-1})R(\varepsilon_t)^-\varepsilon_t.$

则 $x\in\mathscr{H}_t^k$, 且由引理 3.1.3, 对任意 $s\in N_{t-1}$ 有

$$\begin{aligned}
R(x,\ z_s) &= F_s + (F_t - S(F,\ Z^{z_t})_{t-1})R(\varepsilon_t)^- R(\varepsilon_t,\ z_s) = F_s,\\
R(x, z_t) &= R(z(F)_{t-1}, z_t) + (F_t - S(F, Z^{z_t})_{t-1})R(\varepsilon_t)^- R(\varepsilon_t,\ z_t)\\
&= R(z(F)_{t-1},\ \varepsilon_t + z(Z^{z_t})_{t-1}) + C_t R(\varepsilon_t)R(\varepsilon_t)^- R(\varepsilon_t)\\
&= S(F,\ Z^{z_t})_{t-1} + (F_t - S(F,\ Z^{z_t})_{t-1}) = F_t.
\end{aligned}$$

所以 $F\in\mathscr{R}_t^k$ 且 $z(F)_t = x$.

最后，设 $F\in\mathscr{R}_t^k$, $G\in\mathscr{R}_t^l$, 则由已证明部分及引理 3.1.3 知

$$\begin{aligned}
S(F,\ G)_t &= R(z(F)_t,\ z(G)_t)\\
&= R(z(F)_{t-1},\ z(G)_{t-1}) + (F_t - S(F,\ Z^{z_t})_{t-1})\\
&\quad \times R(\varepsilon_t)^- R(\varepsilon_t) R(\varepsilon_t)^- (G_t - S(G,\ Z^{z_t})_{t-1})^*\\
&= S(F,\ G)_{t-1} + (F_t - S(F,\ Z^{z_t})_{t-1})
\end{aligned}$$

$$\times R(\varepsilon_t)^{-}(G_t - S(G, Z^{z_t})_{t-1})^{*}. \blacksquare$$

定理 3.1.4 通过新息 ε_t 及其方差阵 $R(\varepsilon_t)$ 递推地表出了 $z(F)_t$ 和 $S(F, G)_t$（注意 $S(F)_t \triangleq S(F, F)_t$），并给出了 $F \in \mathscr{R}_t^z$ 的递推判别准则（注意：如果对一切 $t \in N$，$R(\varepsilon_t)$ 是可逆的，则显然有任何 $F \in \mathscr{R}_t^z$）。把这些结果代入第二章各种线性估计问题的表达式中，就在原则上统一解决了基于任意有限步观测（即任意有限的 t）的线性统计计算问题. 但实际上(3.1.3)、(3.1.4)式的形式递推在计算量上存在发散的困难. 这正是下一节要讨论和解决的问题.

§3.2 协方差可分解情形

虽然定理 3.1.4 给出了一般情形的递推表达式，但它实际上并不是严格意义下的递推方程. 因为仅根据由该式每一步算出的诸值及下一步的过程值与过程的协方差，并不就能由此递推算出下一步的诸值. 例如，从 t 步到 $t+1$ 步，首先必须算出

$$\varepsilon_{t+1} = z_{t+1} - z(Z^{z_{t+1}})_t \quad 与 \quad R(\varepsilon_{t+1}) = R(z_{t+1}) - S(Z^{z_{t+1}})_t,$$

而为求其中的 $z(Z^{z_{t+1}})_t$ 与 $S(Z^{z_{t+1}})_t$，由(3.1.3)与(3.1.4)式，仅根据 ε_t、$R(\varepsilon_t)$、z_{t+1} 与 $R(z_{t+1}, z_t)$ 还是不够的，还必须事先求出 $z(Z^{z_{t+1}})_{t-1}$，$S(Z^{z_{t+1}})_{t-1}$ 以及 $S(Z^{z_{t+1}}, Z^{z_t})_{t-1}$ 等. 因此，每一步的计算量并不是值定或有界的，它将随着 t 的增长趋于无穷. 为了克服这一困难，必须对 z_N 的协方差函数的形式给予某种本质的限制.

定义 3.2.1 设 z_N 为一 m 维局部正则序列. $\Phi \triangleq \{\Phi_{t,t-1}, t \in N\}$ 为一 n 阶阵列（通常称为*转移阵*）. 按系统理论对转移阵的习惯用法，记

$$\Phi_{ts} \triangleq \begin{cases} I_n, & s = t; \\ \Phi_{t,t-1} \cdots \Phi_{s+1,s}, & (s < t); \\ \Phi_{st}^{-1}, & s > t \text{（若 } \Phi_{t,t-1} \text{ 值为可逆阵. 否则无意义）.} \end{cases} \tag{3.2.1}$$

又对非负整数 q，记

$$(\Phi_t^{(q)})_s \triangleq \Phi_{t,s+q}, \quad s\in N_{t-q}, \quad t>q. \tag{3.2.2}$$

(注意 $\Phi_t^{(s)}$ 是局限于 N_{t-q} 上的 $C^{n\times n}$ 值函数). 若存在 q 以及 N 上的 $C^{m\times n}$ 值函数 $A=(A_s, \ s\in N)$ 与 N 上的 $C^{n\times m}$ 值函数 $B=(B_s, \ s\in N)$, 使对任意 $t>q$ 有 $\Phi_t^{(q)}B\in \mathscr{H}_{t-q}^{n}$, 且有下式成立:

$$R(z_t, \ z_s)=A_t\Phi_{t,s+q}B_s=A_t(\Phi_t^{(q)}B)_s, \quad s\in N_{t-q-1}, \tag{3.2.3}$$

则称 z_N 对 Φ 有 q 步后可分解的协方差. 此处及今后, 当 $t\leqslant 0$ 时, N_t 恒理解为空集.

当 $\Phi_{t,t-1}\equiv I_n$ 时, 这一定义是作者与杜金观[11]首次提出的. $q=0$ 的情形则可见[3]. 又当 $\Phi_{t,t-1}$ 恒为可逆阵时, (3.2.3)式可表成

$$B(z_t, \ z_s)=(A_t\Phi_{t0})(\Phi_{0,s+q}B_s), \ (s\in N_{t-q-1}),$$

即可转化为 $\Phi_{t,t-1}\equiv I_n$ 的情形.

这一定义的用途和它的内含, 通过下面的定理 3.2.3 和定理 3.2.7, 将会变得十分清楚.

引理 3.2.2 设 z_N 对 Φ 有 q 步后可分解的协方差 (3.2.3), 则对任意的 $s\in N_{t-q}$ 及 $F\in \mathscr{H}_s^{n}$ 有下式成立:

$$\left.\begin{aligned} z(\Phi_t^{(q)}B)_s&=\Phi_{t,s+q}z(\Phi_{s+q}^{(q)}B)_s,\\ S(\Phi_t^{(q)}B, \ F)_s&=\Phi_{t,s+q}S(\Phi_{s+q}^{(q)}B, \ F)_s. \end{aligned}\right\} \tag{3.2.4}$$

对任意的 $s\in N_{t-q-1}$ 有下式成立:

$$z(Z^{z_t})_s=A_tz(\Phi_t^{(q)}B)_s, \ R(z_t, \ z(F)_s)=A_tS(\Phi_t^{(q)}B, \ F)_s \tag{3.2.5}$$

证 首先, 对任意 $s\in N_{t-q}$ 及 $r\in N_s$ 有

$$\begin{aligned} &R(\Phi_{t,s+q}z(\Phi_{s+q}^{(q)}B)_s, \ z_r)\\ &=\Phi_{t,s+q}\Phi_{s+q,r+q}B_r=\Phi_{t,r+q}B_r=R(z(\Phi_t^{(q)}B)_s, \ z_r). \end{aligned}$$

故得

$$z(\Phi_t^{(q)}B)_s=\Phi_{t,s+q}z(\Phi_{s+q}^{(q)}B)_s.$$

于是

$$\begin{aligned} S(\Phi_t^{(q)}B, \ F)_s&=R(z(\Phi_t^{(q)}B)_s, \ z(F)_s)\\ &=\Phi_{t,s+q}S(\Phi_{s+q}^{(q)}B, \ F)_s. \end{aligned}$$

又由(3.2.3), 对任意的 $s\in N_{t-q-1}$ 及 $r\in N_s$ 有

$$R(A_tz(\Phi_t^{(q)}B)_s, \ z_r)=A_t\Phi_{t,r+q}B_r=R(z_t, \ z_r)=Z_r^{z_t}.$$

所以 $z(Z^{z_t})_s = A_t z(\Phi_t^{(q)}B)_s$. 由此即推得

$$\begin{aligned}
&R(z_t,\ z(F)_s)\\
&\quad = R(z(Z^{z_t})_s,\ z(F)_s) - A_t R(z(\Phi_t^{(q)}B)_s,\ z(F)_s)\\
&\quad = A_t S(\Phi_t^{(q)}B,\ F)_s. \ \blacksquare
\end{aligned}$$

定理3.2.3 设 z_N 对 Φ 有 q 步后可分解的协方差(3.2.3). 记

$$y_t \triangleq z(\Phi_t^{(q)}B)_{t-q},\quad \widetilde{B}_t \triangleq R(y_{t+q},\ \varepsilon_t),$$

$$J_{ts} \triangleq R(z_t,\ \varepsilon_s),\ (s=t-q,\ \cdots,\ t-1),$$

则 z_N 的新息序列 ε_N 满足如下的递推方程:

$$\begin{aligned}
J_{ts} = &R(z_t,\ z_s) - A_t\Phi_{t,s-1}R(y_{s-1})\Phi^*_{s,s-1}A^*_s\\
&- A_t\sum_{r=s-q}^{t-q-1}\Phi_{t,r+q}\widetilde{B}_r R(\varepsilon_r)^- J^*_{sr} - \sum_{r=t-q}^{s-1}J_{tr}R(\varepsilon_r)^- J^*_{sr},\\
&s=t-q,\ \cdots,\ t-1 \quad (q\geqslant 1),
\end{aligned} \tag{3.2.6}$$

$$\varepsilon_t = z_t - A_t\Phi_{t,t-1}y_{t-1} - \sum_{s=t-q}^{t-1}J_{ts}R(\varepsilon_s)^-\varepsilon_s, \tag{3.2.7}$$

$$R(\varepsilon_t) = R(z_t) - A_t\Phi_{t,t-1}R(y_{t-1})\Phi^*_{t,t-1}A^*_t - \sum_{s=t-q}^{t-1}J_{ts}R(\varepsilon_s)^- J^*_{ts}, \tag{3.2.8}$$

$$\begin{aligned}
\widetilde{B}_{t-q} = &B_{t-q} - \Phi_{t,t-q-1}R(y_{t-q-1})\Phi^*_{t-q,t-q-1}A^*_{t-q}\\
&- \sum_{s=t-q}^{t-1}\Phi_{ts}\widetilde{B}_{s-q}R(\varepsilon_{s-q})^- J^*_{t-q,s-q},
\end{aligned} \tag{3.2.9}$$

$$y_t = \Phi_{t,t-1}y_{t-1} + \widetilde{B}_{t-q}R(\varepsilon_{t-q})^-\varepsilon_{t-q}, \tag{3.2.10}$$

$$R(y_t) = \Phi_{t,t-1}R(y_{t-1})\Phi^*_{t,t-1} + \widetilde{B}_{t-q}R(\varepsilon_{t-q})^-\widetilde{B}^*_{t-q}. \tag{3.2.11}$$

又 $F\in\mathscr{R}_t^q$ 的必要充分条件是: 对任意 $s\in N_t$, 存在 $C_s\in C^{k\times m}$, 使 $\widetilde{F}_s = C_s R(\varepsilon_s)$, 其中 $\widetilde{F}_t$ 按下式递推计算:

$$\begin{aligned}
\widetilde{F}_t &\triangleq F_t - S(F,\ \Phi_{t-1}^{(q)}B)_{t-1-q}\Phi^*_{t,t-1}A^*_t - \sum_{s=t-q}^{t-1}\widetilde{F}_s R(\varepsilon_s)^- J^*_{ts}\\
&= R(z(F)_t,\ \varepsilon_t),
\end{aligned} \tag{3.2.12}$$

$$S(F,\ \Phi_t^{(q)}B)_{t-q} = S(F,\ \Phi_{t-1}^{(q)}B)_{t-1-q}\Phi^*_{t,t-1} + \widetilde{F}_{t-q}R(\varepsilon_{t-q})^-\widetilde{B}^*_{t-q}. \tag{3.2.13}$$

此时 $z(F)_t$、$S(F)_t$ 满足如下递推方程：

$$z(F)_t = z(F)_{t-1} + \widetilde{F}_t R(\varepsilon_t)^- \varepsilon_t, \tag{3.2.14}$$

$$S(F)_t = S(F)_{t-1} + \widetilde{F}_t R(\varepsilon_t)^- \widetilde{F}_t^*. \tag{3.2.15}$$

注意：由协方差可分解的定义，$\widetilde{B}$ 须满足与 $\widetilde{F}$ 相同的条件（取 $h=n$）. 又此处及今后，凡式中出现求和号"Σ"，而其下限超过上限时，其和即约定为 0. 例如(3.2.6)的后一和号当 $s=t-q$ 时，以及(3.2.7)～(3.2.9)、(3.2.12)中当 $q=0$ 时，就是这种情况.

证 由定理 3.1.4 与引理 3.2.2：

$$\begin{aligned}\varepsilon_t &= z_t - z(Z^{z_t})_{t-1}\\ &= z_t - z(Z^{z_t})_{t-q-1} - \sum_{s=t-q}^{t-1} R(z_t, \varepsilon_s) R(\varepsilon_s)^- \varepsilon_s\\ &= z_t - A_t \Phi_{t,t-1} z(\Phi_{t-1}^{(q)} B)_{t-1-q}\\ &\quad - \sum_{s=t-q}^{t-1} J_{ts} R(\varepsilon_s)^- \varepsilon_s.\end{aligned}$$

即证得(3.2.7). 由此及引理 3.1.3 易推出(3.2.8). 同理可得(3.2.10)、(3.2.11)及(3.2.13)～(3.2.15).

由(3.2.7)与引理 3.2.2：

$$\begin{aligned}\widetilde{B}_{t-q} &= R(z(\Phi_t^{(q)} B)_{t-q}, \varepsilon_{t-q})\\ &= B_{t-q} - R(y_t, y_{t-q-1}) \Phi_{t-q,t-q-1}^* A_{t-q}^*\\ &\quad - \sum_{s=t-q}^{t-1} R(y_t, \varepsilon_{s-q}) R(\varepsilon_{s-q})^- J_{t-q,s-q}^*\\ &= B_{t-q} - \Phi_{t,t-q-1} R(y_{t-q-1}) \Phi_{t-q,t-q-1}^* A_{t-q}^*\\ &\quad - \sum_{s=t-q}^{t-1} \Phi_{ts} R(y_s, \varepsilon_{s-q}) R(\varepsilon_{s-q})^- J_{t-q,s-q}^*,\end{aligned}$$

这就是(3.2.9). 类似地可证(3.2.12)的后一等号. 又对任意的 $s\in N_{t-q-1}$,

$$\begin{aligned}R(z_t, \varepsilon_s) &= R(z(Z^{z_t})_s, \varepsilon_s) = A_t R(z(\Phi_t^{(q)} B)_s, \varepsilon_s)\\ &= A_t \Phi_{t,s+q} \widetilde{B}_s.\end{aligned}$$

再由(3.2.7)，对 $s=t-q, \cdots, t-1$,

$$J_{ts} = R(z_t, z_s) - R(z(Z^{z_t})_{s-q-1}, y_{s-1}) \Phi_{s,s-1}^* A_s^*$$

$$-\sum_{r=s-q}^{s-1} R(z_t, \varepsilon_r)R(\varepsilon_r)^- J_{sr}$$

$$=R(z_t, z_s)-A_t\Phi_{t,s-1}R(y_{s-1})\Phi_{s,s-1}^* A_s^*$$

$$-\sum_{r=s-q}^{t-q-1} A_t\Phi_{t,r+q}\widetilde{B}_r R(\varepsilon_r)^- J_{sr}^* - \sum_{r=t-q}^{s-1} J_{tr}R(\varepsilon_r)^- J_{sr}^*.$$

即得(3.2.6).

最后，$F\in\mathscr{R}_t^*$ 的必要充分条件由定理 3.1.4 直接得出. ▌

定理 3.2.3 对 q 步后可分解协方差的随机序列 z_N 给出了其新息序列 ε_N 及逆再生表示的一套完整的递推方程. 由该定理可以看出，只要在 $t-1$ 步保存了已經算出的 q^2+6q+5 组数据：$\{J_{sr}: s-q\leqslant r\leqslant s-1,\ t-q\leqslant s\leqslant t-1\}$，$\{R(\varepsilon_s): t-2q\leqslant s\leqslant t-1\}$，$\{R(y_s): t-q-1\leqslant s\leqslant t-1\}$，$\{\widetilde{B}_{s-q}, \varepsilon_s, \widetilde{F}_s: t-q\leqslant s\leqslant t-1\}$，$y_{t-1}$，$S(F, \Phi_{t-1}^{(q)}B)_{t-1-q}$，$z(F)_{t-1}$，$S(F)_{t-1}$，即可在测得 z_t 后在 t 步依序算出 $q+9$ 组新的数据：$\{J_{ts}: t-q\leqslant s\leqslant t-1\}$，$\varepsilon_t$，$R(\varepsilon_t)$，$\widetilde{B}_{t-q}$，$y_t$，$R(y_t)$，$\widetilde{F}_t$，$S(F, \Phi_t^{(q)}B)_{t-q}$，$z(F)_t$，$S(F)_t$，并能判断是否 $F\in\mathscr{R}_t^*$. 按定理 3.1.4 后面的说明，这就统一解决了基于对 z_t 的任意有限步观测的线性统计递推计算问题.

对可分解协方差的某些特殊情形，递推方程还可大大简化.

系 3.2.4　当 $q=1$ 或 $\Phi_{t,t-1}$ 恒为可逆阵时，定理 3.2.3 中的(3.2.6)可简化为

$$J_{ts}=R(z_t, z_s)-A_t\Phi_{t,s+q}(R_s-\widetilde{B}_s)$$

$$-\sum_{r=s-q}^{s-1}(J_{tr}-A_t\Phi_{t,r+q}\widetilde{B}_r)R(\varepsilon_r)^- J_{sr}^*,$$

$$s=t-q, \cdots, t-1\ (q\geqslant 1). \tag{3.2.6'}$$

而(3.2.9)可改换为

$$\widetilde{B}_t=B_t-\Phi_{t+q,t-1}R(y_{t-1})\Phi_{t,t-1}^* A_t^*-\sum_{s=t-q}^{t-1}\Phi_{t+q,s+q}\widetilde{B}_s R(\varepsilon_s)^- J_{ts}^*. \tag{3.2.9'}$$

证　(3.2.9′)就是(3.2.9)用 t 代替 $t-q$ 的结果，至于(3.2.6′)，由(3.2.9′)及(3.2.1)规定的记法，对 $s=t-q, \cdots, t-1$，

$$\Phi_{t,s+q}(B_s-\widetilde{B}_s)=\Phi_{t,s-1}R(y_{s-1})\Phi^*_{s,s-1}A^*_s$$
$$+\sum_{r=s-q}^{s-1}\Phi_{t,r+q}\widetilde{B}_rR(\varepsilon_r)^-J^*_{sr},$$

再由(3.2.6)即有

$$J_{ts}=R(z_t,\ z_s)-A_t(\Phi_{t,s-1}R(y_{s-1})\Phi^*_{s,s-1}A^*_s$$
$$+\sum_{r=s-q}^{t-q-1}\Phi_{t,r+q}\widetilde{B}_rR(\varepsilon_r)^-J^*_{sr})$$
$$-\sum_{r=t-q}^{s-1}J_{tr}R(\varepsilon_r)^-J^*_{sr}$$
$$=R(z_t,\ z_s)-A_t\Phi_{t,s+q}(B_s-\widetilde{B}_s)$$
$$-\sum_{r=t-q}^{s-1}(J_{tr}-A_t\Phi_{t,r+q}\widetilde{B}_r)R(\varepsilon_r)^-J^*_{sr}. \blacksquare$$

系 3.2.5 设 z_N 有 0 步后可分解的协方差，即可表为

$$R(z_t,\ z_s)=A_t\Phi_{ts}B_s,\ (t\in N,\ s\in N_{t-1}). \qquad (3.2.16)$$

则简单地有如下递推方程成立：

$$\varepsilon_t=z_t-A_t\Phi_{t,t-1}y_{t-1},$$
$$R(\varepsilon_t)=R(z_t)-A_t\Phi_{t,t-1}R(y_{t-1})\Phi^*_{t,t-1}A^*_t,$$
$$y_t=\Phi_{t,t-1}y_{t-1}+\widetilde{B}_tR(\varepsilon_t)^-\varepsilon_t,$$
$$R(y_t)=\Phi_{t,t-1}R(y_{t-1})\Phi^*_{t,t-1}+\widetilde{B}_tR(\varepsilon_t)^-\widetilde{B}^*_t,$$

其中 $\widetilde{B}_t\triangleq B_t-\Phi_{t,t-1}R(y_{t-1})\Phi^*_{t,t-1}A^*_t$. 又 $F\in\mathscr{R}^*_t$ 的必要充分条件是：$F\in\mathscr{R}^*_{t-1}$ 且存在 $C_t\in C^{k\times m}$ 使

$$\widetilde{F}_t\triangleq F_t-S(F,\ \Phi^{(0)}_{t-1}B)_{t-1}\Phi^*_{t,t-1}A^*_t=C_tR(\varepsilon_t).$$

此时有以下式子成立：

$$z(F)_t=z(F)_{t-1}+\widetilde{F}_tR(\varepsilon_t)^-\varepsilon_t,$$
$$S(F)_t=S(F)_{t-1}+\widetilde{F}_tR(\varepsilon_t)^-\widetilde{F}^*_t,$$
$$S(F,\ \Phi^{(0)}_tB)_t=S(F,\ \Phi^{(0)}_{t-1}B)_{t-1}\Phi^*_{t,t-1}+\widetilde{F}_tR(\varepsilon_t)^-\widetilde{B}^*_t.$$

证 只须注意当 $q=0$ 时定理 3.2.3 中关于求和号的约定，即证得所有结论. ∎

系 3.2.6 设 z_N 在 q 步后(严格大于 q 步)不相关，即有

$$R(z_t,\ z_s)=0,\ |t-s|>q. \qquad (3.2.17)$$

则简单地有如下递推方程成立：

$$J_{ts}=R(z_t,\ z_s)-\sum_{r=t-q}^{s-1}J_{tr}R(\varepsilon_r)^-J_{sr}^*,\ (s=t-q,\ \cdots,\ t-1),$$

$$\varepsilon_t=z_t-\sum_{s=t-q}^{t-1}J_{ts}R(\varepsilon_s)^-\varepsilon_s,$$

$$R(\varepsilon_t)=R(z_t)-\sum_{s=t-q}^{t-1}J_{ts}R(\varepsilon_s)^-J_{ts}^*.$$

$F\in\mathscr{R}_t^q$ 的必要充分条件是：对任何 $s\in N_t$，存在 $C_s\in C^{k\times m}$ 使 $\tilde{F}_s=C_sR(\varepsilon_s)$，其中 $\tilde{F}_t$ 按下式递推：

$$\tilde{F}_t=F_t-\sum_{s=t-q}^{t-1}\tilde{F}_sR(\varepsilon_s)^-J_{ts}^*.$$

此时有下式成立：

$$z(F)_t=z(F)_{t-1}+\tilde{F}_tR(\varepsilon_t)^-\varepsilon_t,$$

$$S(F)_t=S(F)_{t-1}+\tilde{F}_tR(\varepsilon_t)^-\tilde{F}_t^*.$$

证 在定理3.2.3中置 $A=0_{m\times n}$，$B=0_{n\times m}$ 即得. ∎

下一定理是一个基本的表现定理，它表明了 q 步后可分解协方差的随机序列等价于一类非白噪声线性系统的输出(或量测)序列. 它揭示了定义3.2.1的内含，在理论和应用上都有重要意义.

定理3.2.7 设 z_N 是一 m 维二阶随机序列，$\boldsymbol{\Phi}=\{\Phi_{t,t-1},\ t\in N\}$ 为一 n 阶阵列. 则 z_N 对 $\boldsymbol{\Phi}$ 有 q 步后可分解的协方差(3.2.3)的必要充分条件是：z_N 为如下线性系统的输出：

$$\begin{cases}x_t=\Phi_{t,t-1}x_{t-1}+u_t;\\ z_t=A_tx_t+v_t.\end{cases}\tag{3.2.18}$$

其中 $\begin{pmatrix}u_N\\ v_N\end{pmatrix}$ 是 q 步后(严格大于 q 步)不相关的 $n+m$ 维局部正则序列.

证 必要性：设 Φ_{ts}、$\Phi_t^{(q)}$、A、B 分别由(3.2.1)、(3.2.2)、(3.2.3)给出，y_t、$\tilde{B}_t$、J_{ts} 由定理3.2.3给出. 令 $x_t=y_t$，

$$u_t=\tilde{B}_{t-q}R(\varepsilon_{t-q})^-\varepsilon_{t-q},\quad v_t=\varepsilon_t+\sum_{s=t-q}^{t-1}J_{ts}R(\varepsilon_s)^-\varepsilon_s-A_tu_t,$$

则由(3.2.10)、(3.2.7)即得(3.2.18)，且由 z_N 的局部正则性及引理3.1.3知 u_N、v_N 满足定理要求的条件.

充分性：设 z_N 为(3.2.18)的输出. 则对任意 $s\in N_t$,

$$x_t=\Phi_{ts}x_s+\sum_{r=s+1}^{t}\Phi_{tr}u_r.$$

特别有

$$z_t=A_t\sum_{s=1}^{t}\Phi_{ts}u_s+v_t.$$

于是由 $\begin{pmatrix}u_N\\ v_N\end{pmatrix}$ 的局部正则性及 q 步后不相关性，从引理 3.1.3 推出 z_N 的局部正则性，且对任意的 $s\in N_{t-q-1}$ 有

$$\begin{aligned}R(z_t,\ z_s)&=R\left(A_t\left[\Phi_{t,s+q}x_{s+q}+\sum_{r=s+q+1}^{t}\Phi_{tr}u_r\right]+v_t,\ z_s\right)\\&=A_t\Phi_{t,s+q}R(x_{s+q},\ z_s).\end{aligned}\tag{3.2.19}$$

令 $B_s\triangleq R(x_{s+q},\ z_s)$，并注意对任意的 $t>q$、$s\in N_{t-q}$ 有

$$\begin{aligned}(\Phi_t^{(q)}B)_s&=\Phi_{t,s+q}R(x_{s+q},\ z_s)=R(\Phi_{t,s+q}x_{s+q},\ z_s)\\&=R\left(x_t-\sum_{r=s+q+1}^{t}\Phi_{tr}u_r,\ z_s\right)=R(x_t,\ z_s).\end{aligned}\tag{3.2.20}$$

成立.

即知 $\Phi_t^{(q)}B\in\mathscr{R}_{t-q}^{n}$. 根据定义 3.2.1, z_N 有 q 步后可分解的协方差. ▌

由定理 3.2.7 可知，在滤波与控制理论中常见的白噪声线控系统 $\left(\text{即}\begin{pmatrix}u_N\\ v_N\end{pmatrix}\text{为 0 步后不相关(正交)随机序列}\right)$，其输出(或量测) z_N 是 0 步后可分解协方差的随机序列(3.2.16).

§ 3.3　随机序列的滑动和与时变 ARMA 序列

设 z_N 为一 m 维随机序列, $p\in N$. 又对任意的 $t\in N$, $j=1,\cdots,p$, 设 $\Psi_j(t)\in C^{m\times m}$. 令

$$\bar{z}_t\triangleq z_t-\sum_{j=1}^{p}\Psi_j(t)z_{t-j},\quad(t\in N).\tag{3.3.1}$$

这里按通常约定，对非正指标 $s\leqslant 0$, $z_s\triangleq 0$. $\bar{z}_N$ 称为 z_N 的 p 阶滑动和.

定理3.3.1 z_N 为局部正则序列，当且仅当 $\bar{z}_N$ 为局部正则序列．此时，对任意 $t\in N$，$\mathscr{H}(\bar{z},t)=\mathscr{H}(z,t)$，$\bar{z}_N$ 的新息序列 $\bar{\varepsilon}_N$ 等于 z_N 的新息序列 ε_N．又对任一 $F\in\mathscr{R}(z,t)^k$，令

$$\bar{F}_s\triangleq F_s-\sum_{j=1}^{p}F_{s-j}\Psi_j^*(s),\ (s\in N_t),\qquad(3.3.2)$$

则 $F\mapsto\bar{F}$ 为 $R(z,t)^k\to R(\bar{z},t)^k$ 上的一一映射，且

$$\bar{z}(\bar{F})_t=z(F)_t,\ S_{\bar{z}}(\bar{F})_t=S_z(F)_t.$$

证 设 z_N 为局部正则序列．由滑动和的定义，显然有 $\bar{z}_t\in\mathscr{H}(z,t)^m$，故 $\mathscr{H}(\bar{z},t)\subset\mathscr{H}(z,t)$，且由引理 3.1.3 知 $\bar{z}_N$ 为局部正则的．反之，设 $\bar{z}_N$ 是局部正则的．由 $\bar{z}_1=z_1$ 知 $\mathscr{H}(\bar{z},1)=\mathscr{H}(z,1)$．若已证 $\mathscr{H}(\bar{z},t-1)\supset\mathscr{H}(z,t-1)$，则由

$$z_t=\bar{z}_t+\sum_{j=1}^{p}\Psi_j(t)z_{t-j}\in\mathscr{H}(\bar{z},t)^m$$

即推出 $\mathscr{H}(\bar{z},t)\supset\mathscr{H}(z,t)$ 以及 z_N 的局部正则性．

又由新息序列的正交性，对任意的 $s\in N_{t-1}$，$R(\bar{z}_t-\varepsilon_t,\bar{z}_s)=R(\bar{z}_t,\bar{z}_s)$，且

$$\begin{aligned}\bar{z}_t-\varepsilon_t&=z(Z^{z_t})_{t-1}-\sum_{j=1}^{p}\Psi_j(t)z_{t-j}\in\mathscr{H}(z,t-1)^m\\&=\mathscr{H}(\bar{z},t-1)^m,\end{aligned}$$

所以 $\bar{\varepsilon}_t=\varepsilon_t$．

现在，设 $F\in\mathscr{R}(z,t)^k$，则对任意的 $s\in N_t$，

$$\begin{aligned}\bar{F}_s&=F_s-\sum_{j=1}^{p}F_{s-j}\Psi_j^*(s)=R\Big(z(F)_t,\ z_s-\sum_{j=1}^{p}\Psi_j(s)z_{s-j}\Big)\\&=R(z(F)_t,\bar{z}_s).\end{aligned}$$

且 $z(F)_t\in\mathscr{H}(z,t)^k=\mathscr{H}(\bar{z},t)^k$．所以 $\bar{F}\in\mathscr{R}(\bar{z},t)^k$，且 $\bar{z}(\bar{F})_t=z(F)_t$，$S_{\bar{z}}(\bar{F})_t=R(\bar{z}(\bar{F})_t)=R(z(F)_t)=S_z(F)_t$．反之，设 $\bar{F}\in\mathscr{R}(\bar{z},t)^k$，令 $F_s\triangleq R(\bar{z}(\bar{F})_t,z_s)$，$s\in N_t$．则 $F\in\mathscr{R}(z,t)^k$，且

$$\begin{aligned}\bar{F}_s&=R(\bar{z}(\bar{F})_t,\bar{z}_s)=R\Big(\bar{z}(\bar{F})_t,\ z_s-\sum_{j=1}^{p}\Psi_j(s)z_{s-j}\Big)\\&=F_s-\sum_{j=1}^{p}F_{s-j}\Psi_j^*(s).\ \blacksquare\end{aligned}$$

定义 3.3.2　设 z_N 为一 m 维局部正则序列．若存在 $p\in N$、非负整数 q 及 $\Psi_j(t)\in C^{m\times m}$，$(j=1,\cdots,p)$，使

$$\bar{z}_t=z_t-\sum_{j=1}^{p}\Psi_j(t)z_{t-j},\ t\in N,$$

且满足

$$R(\bar{z}_t,\ \varepsilon_s)=0,\ (s\in N_{t-q-1},\ t\in N\backslash N_p),$$

则称 z_N 为时变自回归滑动和序列，或简称时变 ARMA$(p,\ q)$序列．它自然包括通常的(时不变)ARMA$(p,\ q)$序列．

定理 3.3.3　设 z_N 为一 m 维时变 ARMA$(p,\ q)$序列．则当 $p\leqslant q+1$ 时，其新息序列 ε_N 满足如下递推方程：

$$J_{ts}=R(\bar{z}_t,\ \bar{\varepsilon}_s)=R(\bar{z}_t,\ \bar{z}_s)-\sum_{r=t-q}^{s-1}J_{tr}R(\varepsilon_r)^{-}J_{sr}^{*},\ (s=t-q,\ \cdots,\ t-1),\tag{3.3.3}$$

$$\varepsilon_t=z_t-\sum_{j=1}^{p}\Psi_j(t)z_{t-j}-\sum_{s=t-q}^{t-1}J_{ts}R(\varepsilon_s)^{-}\varepsilon_s,\tag{3.3.4}$$

$$R(\varepsilon_t)=R(\bar{z}_t)-\sum_{s=t-q}^{t-1}J_{ts}R(\varepsilon_s)^{-}J_{ts}^{*}.\tag{3.3.5}$$

又 $F\in\mathscr{R}(z,\ t)^k$ 的必要充分条件是 $F\in\mathscr{R}(z,\ t-1)^k$ 且存在 $C_t\in C^{k\times m}$ 使 $\tilde{F}_t=C_tR(\varepsilon_t)$，其中递推地定义

$$\tilde{F}_t\triangleq F_t-\sum_{j=1}^{p}F_{t-j}\Psi_j^{*}(t)-\sum_{s=t-q}^{t-1}\tilde{F}_sR(\varepsilon_s)^{-}J_{ts}^{*},\tag{3.3.6}$$

此时有下式成立：

$$z(F)_t=z(F)_{t-1}+\tilde{F}_tR(\varepsilon_t)^{-}\varepsilon_t,\tag{3.3.7}$$

$$S(F)_t=S(F)_{t-1}+\tilde{F}_tR(\varepsilon_t)^{-}\tilde{F}_t^{*}.\tag{3.3.8}$$

当 $p>q+1$ 时，则只须对 $t\in N_p$ 将前 4 个递推式的和号下标 $r=t-q$、$s=t-q$ 分别修改为 $r=1$、$s=1$ 即可(对 $t\in N\backslash N_p$ 则不作修改)．

证　由定理 3.3.1 与系 3.2.6 即得．∎

第 4 章

离散时间系统的线性滤波

记号 4.0.1 在本章中，我们仍沿用前三章，尤其是第三章中规定的记号．设 z_N 为一 m 维局部正则序列，x_N 为一 n 维二阶随机序列．我们的目的是要通过量测 z_N 来对 x_N 作线性估计．具体地说，设 τ、$t\in N$，用 $\hat{x}_{\tau|t}$ 表基于到 t 为止的量测 z_{N_t} 对 x_τ 所作的线性最小方差预报（参见第二章 § 2.1）：

$$\hat{x}_{\tau|t} \triangleq \pi(x_\tau | z_{N_t}),$$

且记 $P_{\tau|t} \triangleq R(x_\tau - \hat{x}_{\tau|t})$．根据第二章的结果，$\hat{x}_{\tau|t}$ 与 $P_{\tau|t}$ 完全由 z_{N_t} 的再生核表示空间 $(\mathscr{R}_t, S(\cdot)_t)$ 及逆再生表示 $z(\cdot)_t$ 确定．如果进一步假定 z_N 有 q 步后可分解的协方差，则利用第三章 § 2 的结果，$\hat{x}_{\tau|t}$ 和 $P_{\tau|t}$ 可以按 t 进行递推计算（固定 τ）．但是，在实际应用，尤其是有关动态数据适时处理及控制系统的应用中，为了大大减少数据存贮量和快速计算的需要，还须给出当 τ 与 t 一道增长（$\tau-t$ 固定）时的递推算法．这就是本章要讨论的主要问题，通常称为离散时间系统的线性滤波问题．在工程应用中，对 $\tau=t$、$\tau>t$、$\tau<t$ 三种情形，$\hat{x}_{\tau|t}$ 分别称为 x_τ 的滤波（即 $\hat{x}_{t|t}$），$\tau-t$ 步预测（又称外推），$t-\tau$ 步平滑（又称内插）．

§ 4.1 一类非白噪声线性系统的滤波

在定理 3.2.7 中已经证明，z_N 对 $\Phi=\{\Phi_{t,t-1},\ t\in N\}$ 有 q 步后可分解的协方差（3.2.3）的必要充分条件是：z_N 为如下线性系统的输出：

$$
\begin{cases} x_t = \Phi_{t,t-1} x_{t-1} + u_t, \\ z_t = A_t x_t + v_t, \end{cases} \tag{4.1.1}
$$

其中$\begin{pmatrix} u_N \\ v_N \end{pmatrix}$是 q 步后不相关的 $n+m$ 维局部正则序列. 通常称(4.1.1)的前一式为系统的状态(动态)方程, 后一式为量测方程, x_t 称为 t 时的状态, u_N、v_N 分别称为动态噪声与量测噪声。

下面考虑这个系统的滤波问题:

首先假定该系统不含未知参数, 即 $\{\Phi_{t,t-1}\}$、$\{A_t\}$ 以及 u_N、v_N 的均值与协方差函数(包括自协方差、互协方差)都完全已知, 且依约定 $x_0=0$. 这已经大大扩充了熟知的 Kalman 滤波所考虑的白噪声线性系统(即 $q=0$ 的情形).

定理 4.1.1 对于有 q 步后不相关噪声的线性系统 (4.1.1), 记

$$
U_{ts} \triangleq R(u_t, u_s),\ V_{ts} \triangleq R(v_t, v_s),\ W_{ts} \triangleq R(u_t, v_s),\ |t-s| \leqslant q. \tag{4.1.2}
$$

则 z_N 的新息 ε_N 与状态的 q 步预测 $\hat{x}_{t|t-q}$ 及其误差方差阵 $P_{t|t-q}$ 之间有如下的递推关系:

$$
\hat{x}_{t|t-q-1} = \Phi_{t,t-1} \hat{x}_{t-1|t-1-q} + E u_t, \tag{4.1.3}
$$

$$
P_{t|t-q-1} = \Phi_{t,t-1} P_{t-1|t-1-q} \Phi^*_{t,t-1} + \sum_{s=t-q}^{t-1} (\Phi_{ts} U_{st} + U_{ts} \Phi^*_{ts}) + U_{tt}, \tag{4.1.4}
$$

$$
\begin{aligned}
J_{ts} = {} & A_t \Big(\Phi_{ts} P_{s|s-q-1} + \sum_{r=s+1}^{t} \Phi_{tr} \sum_{l=r-q}^{s} U_{rl} \Phi^*_{sl} \Big) A^*_s \\
& + A_t \Big(\sum_{r=s-q}^{t} \Phi_{tr} W_{rs} - \sum_{r=s-q}^{t-q-1} \Phi_{t,r+q} \tilde{B}_r R(\varepsilon_r)^- J^*_{sr} \Big) \\
& + \sum_{r=t-q}^{s} W^*_{rt} \Phi^*_{sr} A^*_s + V_{ts} - \sum_{r=s-q}^{s-1} J_{tr} R(\varepsilon_r)^- J^*_{sr}, \\
& s = t-q, \cdots, t-1 \quad (q \geqslant 1).
\end{aligned} \tag{4.1.5}
$$

$$
\varepsilon_t = z_t - A_t \hat{x}_{t|t-q-1} - \sum_{s=t-q}^{t-1} J_{ts} R(\varepsilon_s)^- \varepsilon_s - E v_t, \tag{4.1.6}
$$

$$R(\varepsilon_t)=A_tP_{t|t-q-1}A_t^*+\sum_{s=t-q}^{t}(A_t\Phi_{ts}W_{st}+W_{st}^*\Phi_{ts}^*A_t^*)$$
$$+V_{tt}-\sum_{s=t-q}^{t-1}J_{ts}R(\varepsilon_s)^-J_{ts}^*, \tag{4.1.7}$$

$$\widetilde{B}_{t-q}=\Big(\Phi_{t,t-q}P_{t-q|t-2q-1}+\sum_{s=t-q+1}^{t}\Phi_{ts}\sum_{r=s}^{t}U_{s,r-q}\Phi_{t-q,r-q}^*\Big)A_{t-q}^*$$
$$+\sum_{s=t-q}^{t+q}\Phi_{t,s-q}W_{s-q,t-q}-\sum_{s=t-q}^{t-1}\Phi_{ts}\widetilde{B}_{s-q}R(\varepsilon_{s-q})^-J_{t-q,s-q}^*, \tag{4.1.8}$$

$$\hat{x}_{t|t-q}=\hat{x}_{t|t-q-1}+\widetilde{B}_{t-q}R(\varepsilon_{t-q})^-\varepsilon_{t-q}, \tag{4.1.9}$$

$$P_{t|t-q}=P_{t|t-q-1}-\widetilde{B}_{t-q}R(\varepsilon_{t-q})^-\widetilde{B}_{t-q}. \tag{4.1.10}$$

此外，当 $q=1$ 或(4.1.1)为可逆系统(即 $\Phi_{t,t-1}$ 恒为可逆阵)时，(4.1.5)可简化为

$$J_{ts}=A_t\Big[\Phi_{t,s+q}\widetilde{B}_s-\sum_{r=t+1}^{s+q}\Phi_{tr}\Big(\sum_{l=r-q}^{s}U_{rl}\Phi_{sl}^*A_s^*+W_{rs}\Big)\Big]$$
$$+\sum_{r=t-q}^{s}W_{rt}^*\Phi_{sr}^*A_s^*+V_{ts}$$
$$-\sum_{r=t-q}^{s-1}(J_{tr}-A_t\Phi_{t,r+q}\widetilde{B}_r)R(\varepsilon_r)^-J_{sr}^*,$$
$$(s=t-q,\ \cdots,\ t-1); \tag{4.1.5'}$$

而(4.1.8)改换成

$$\widetilde{B}_t=\Big(\Phi_{t+q,t}P_{t|t-q-1}+\sum_{s=t+1}^{t+q}\Phi_{t+q,s}\sum_{r=s-q}^{t}U_{sr}\Phi_{tr}^*\Big)A_t^*$$
$$+\sum_{s=t-q}^{t+q}\Phi_{t+q,s}W_{st}-\sum_{s=t-q}^{t-1}\Phi_{t+q,s+q}\widetilde{B}_sR(\varepsilon_s)^-J_{ts}^*. \tag{4.1.8'}$$

证 首先，由于 u_N、v_N 的均值函数都是已知的，在整个证明过程中，不妨设 Eu_t、Ev_t 都恒等于 0，从而 Ex_t、Ez_t 也恒为 0。否则，可换为考虑线性系统

$$\begin{cases}x_t-Ex_t=\Phi_{t,t-1}(x_{t-1}-Ex_{t-1})+u_t-Eu_t;\\ z_t-Ez_t=A_t(x_t-Ex_t)+v_t-Ev_t.\end{cases}$$

在此假定下，由定理 2.1.2 知，对任意的 t、$s\in N$，

$$\hat{x}_{t|s}=z(Z^{x_t})_s,\quad P_{t|s}=R(x_t)-S(Z^{x_t})_s=R(x_t)-R(\hat{x}_{t|s}). \tag{4.1.11}$$

又由 (3.2.19) 与 (3.2.20), 对 $s\in N_{t-q-1}$, $R(z_t,\ z_s)=A_t\Phi_{t,s+q}R(x_{s+q},\ z_s)\triangleq A_t\Phi_{t,s+q}B_s$; 对 $s\in N_{t-q}$, $(\Phi_t^{(q)}R)_s\triangleq\Phi_{t,s+q}B_s=R(x_t,\ z_s)$, 从而 $\Phi_t^{(q)}R\in\mathscr{R}^n_{t-q}$. 由此及引理 3.2.2 知, 对任意 $s\in N_{t-q}$,

$$\hat{x}_{t|s}=z(\Phi_t^{(q)}R)_s=\Phi_{t,s+q}\hat{x}_{s+q|s},\quad R(\hat{x}_{t|s})=\Phi_{t,s+q}R(\hat{x}_{s+q|s})\Phi^*_{t,s+q}. \tag{4.1.12}$$

再由(4.1.1),对 $t\in N$, $t-q\leqslant s\leqslant t$, 容易算出如下诸式:

$$\begin{aligned}R(x_t)&=R(\Phi_{t,t-1}x_{t-1}+u_t)\\&=\Phi_{t,t-1}R(x_{t-1})\Phi^*_{t,t-1}+U_{tt}+R\left(\sum_{s=1}^{t-1}\Phi_{ts}u_s,\ u_t\right)\\&\quad+R\left(u_t,\ \sum_{s=1}^{t-1}\Phi_{ts}u_s\right)\\&=\Phi_{t,t-1}R(x_{t-1})\Phi^*_{t,t-1}+U_{tt}+\sum_{s=t-q}^{t-1}(\Phi_{ts}U_{st}+U_{ts}\Phi^*_{ts}),\end{aligned} \tag{4.1.13}$$

$$\begin{aligned}R(x_t,\ z_s)&=R(x_t,\ A_sx_s+v_s)\\&=R\left(\Phi_{ts}x_s+\sum_{r=s+1}^{t}\Phi_{tr}u_r,\ x_s\right)A_s^*+R(x_t,\ v_s)\\&=\Phi_{ts}R(x_s)A_s^*+\sum_{r=s+1}^{t}\Phi_{tr}R\left(u_r,\ \sum_{l=1}^{s}\Phi_{sl}u_l\right)A_s^*\\&\quad+R\left(\sum_{r=1}^{t}\Phi_{tr}u_r,\ v_s\right)\\&=\Phi_{ts}R(x_s)A_s^*+\sum_{r=s+1}^{t}\Phi_{tr}\sum_{l=r-q}^{s}U_{rl}\Phi^*_{sl}A^*_s\\&\quad+\sum_{r=s-q}^{t}\Phi_{tr}W_{rs},\end{aligned} \tag{4.1.14}$$

$$\begin{aligned}R(z_t,\ z_s)&=R(A_tx_t+v_t,\ z_s)\\&=A_tR(x_t,\ z_s)+R\left(v_t,\ A_s\sum_{r=1}^{s}\Phi_{sr}u_r+v_s\right)\\&=A_tR(x_t,\ z_s)+\sum_{r=t-q}^{s}W^*_{rt}\Phi^*_{sr}A^*_s+V_{ts}.\end{aligned} \tag{4.1.15}$$

现在，在(4.1.12)与(4.1.11)中令 $s=t-q-1$，由(4.1.13)即得(4.1.3)与(4.1.4)。再将以上诸结果代入定理3.2.3中的(3.2.6)～(3.2.11)诸式(其中的 $y_t=\hat{x}_{t|t-q}$, $\tilde{B}_{t-q}=R(x_t, \varepsilon_{t-q})$, $B_{t-q}=R(x_t, z_{t-q})$)，就分别得到(4.1.5)～(4.1.10)。至于 $q=1$ 或可逆系统情形的(4.1.5′)，易从系3.2.4的(3.2.6′)算出。

∎

有了新息的递推公式，就不难求出状态的预测、滤波与平滑的递推方程。

定理 4.1.2 对于线性系统(4.1.1)，设 U、V、W 如(4.1.2)定义，ε_t、$R(\varepsilon_t)$、$\tilde{B}_t$ 由定理4.1.1算出，$\tau=t+\delta$，其中 δ 为一个固定整数。则 $\hat{x}_{\tau|t}$、$P_{\tau|t}$ 的递推公式如下：

① 当 $\delta\geqslant q$ 时，有下式成立：

$$\hat{x}_{\tau|t-1}=\Phi_{\tau,\tau-1}\hat{x}_{\tau-1|t-1}+Eu_\tau, \tag{4.1.16}$$

$$\begin{aligned}\hat{x}_{\tau|t}&=\hat{x}_{\tau|t-1}+\Phi_{\tau,t+q}\tilde{B}_tR(\varepsilon_t)^-\varepsilon_t\\&=\Phi_{\tau,t+q}\hat{x}_{t+q|t}+\sum_{s=t+q+1}^{\tau}Eu_s,\end{aligned} \tag{4.1.17}$$

$$P_{\tau|t-1}=\Phi_{\tau,\tau-1}P_{\tau-1|t-1}\Phi^*_{\tau,\tau-1}+\sum_{s=\tau-q}^{\tau-1}(\Phi_{\tau s}U_{s\tau}+U_{\tau s}\Phi^*_{\tau s})+U_{\tau\tau}, \tag{4.1.18}$$

$$P_{\tau|t}=P_{\tau|t-1}-\Phi_{\tau,t+q}\tilde{B}_tR(\varepsilon_t)^-\tilde{B}^*_t\Phi^*_{\tau,t+q}. \tag{4.1.19}$$

②当 $\delta<q$ 时，在每一时刻 t 还须递推 $q-\delta$ 次：

$$\hat{x}_{\tau|s}=\hat{x}_{\tau|s-1}+B_{\tau s}R(\varepsilon_s)^-\varepsilon_s, \tag{4.1.20}$$

$$P_{\tau|s}=P_{\tau|s-1}-B_{\tau s}R(\varepsilon_s)^-B^*_{\tau s},\ (s=\tau-q+1,\ \cdots,\ t). \tag{4.1.21}$$

其中

$$\begin{aligned}B_{\tau s}=&\Big(\Phi_{\tau s}P_{s|s-q-1}+\sum_{r=s+1}^{\tau}\Phi_{\tau r}\sum_{l=r-q}^{s}U_{rl}\Phi^*_{sl}\Big)A^*_s+\sum_{r=s-q}^{\tau}\Phi_{\tau r}W_{rs}\\&-\sum_{r=s-q}^{\tau-q}\Phi_{\tau,r+q}\tilde{B}_rR(\varepsilon_r)^-J^*_{sr}-\sum_{r=\tau-q+1}^{s-1}B_{\tau r}R(\varepsilon_r)^-J^*_{sr},\\&(s=\tau-q+1,\ \cdots,\ \tau\wedge t),\end{aligned} \tag{4.1.22}$$

$$B_{\tau s}=\Big(P_{\tau\ \tau-q}\Phi^*_{s\tau}+\sum_{r=\tau+1}^{s\wedge(\tau+q)}\sum_{l=r-q}^{\tau}\Phi_{\tau l}U_{lr}\Phi^*_{sr}\Big)A^*_s+\sum_{r=s-q}^{\tau}\Phi_{\tau r}W_{rs}$$

$$-\sum_{r=\tau-q+1}^{s-q-1}B_{\tau r}R(\varepsilon_r)^-\tilde{B}^*_r\Phi^*_{s,r+q}A^*_s-\sum_{r=s-q}^{s-1}B_{\tau r}R(\varepsilon_r)^-J^*_{sr},$$

$$(s=\tau+1,\ \cdots,\ t).\tag{4.1.23}$$

注意最后一式只在考虑平滑问题$(\tau<t)$时才用到. 又式中的$a\wedge b\triangleq\min(a,\ b)$.

证 仍不妨设u_N、v_N的均值为0. 这时,对$\delta\geqslant q$即$\tau\geqslant t+q$的情形,(4.1.16)~(4.1.19)由(4.1.11)~(4.1.13)以及在(4.1.9)与(4.1.10)中以t代替$t-q$立即得出.

对$\delta<q$即$\tau<t+q$的情形,记

$$F^{(\tau)}_s\triangleq R(x_\tau,\ z_s),\ \tilde{F}^{(\tau)}_s=R(x_\tau,\ \varepsilon_s)\triangleq B_{\tau s}.$$

则
$$z(F^{(\tau)})_s=\hat{x}_{\tau|s},\ S(F^{(\tau)})_s=R(\hat{x}_{\tau|s}).$$

于是由(3.2.14)与(3.2.15)即得(4.1.20)与(4.1.21). 又注意当$s\in N_{\tau-q}$时,$B_{\tau s}=\Phi_{\tau,s+q}\tilde{B}_s$. 再由(3.2.12)、(4.1.14),当$s=\tau-q+1,\ \cdots,\ \tau\wedge t$时,

$$B_{\tau s}=R(x_\tau,\ z_s)-R(x_\tau,\ \hat{x}_{s|s-q-1})A^*_s-\sum_{r=s-q}^{s-1}B_{\tau r}R(\varepsilon_r)^-J^*_{sr}$$

$$=R(x_\tau,\ z_s)-\Phi_{\tau s}B(\hat{x}_{s|s-q-1})A^*_s-\sum_{r=s-q}^{s-1}B_{\tau r}R(\varepsilon_r)^-J^*_{sr}$$

$$=\Phi_{\tau s}P_{s|s-q-1}A^*_s+\sum_{r=s+1}^{\tau}\Phi_{\tau r}\sum_{l=r-q}^{s}U_{rl}\Phi^*_{sl}A^*_s+\sum_{r=s-q}^{\tau}\Phi_{\tau r}W_{rs}$$

$$-\sum_{r=s-q}^{\tau-q}\Phi_{\tau,r+q}\tilde{B}_rR(\varepsilon_r)^-J^*_{sr}-\sum_{r=\tau-q+1}^{s-1}B_{\tau r}R(\varepsilon_r)^-J^*_{sr}.$$

此即(4.1.22). 当$\tau<t$且$s=\tau+1,\ \cdots,\ t$时,注意

$$R(x_\tau,\ z_s)=R(x_\tau,\ A_s\Phi_{s\tau}x_\tau)$$

$$+R\Big(\sum_{l=1}^{\tau}\Phi_{\tau l}u_l,\ A_s\sum_{r=\tau+1}^{s}\Phi_{sr}u_r+v_s\Big)$$

$$=R(x_\tau)\Phi^*_{s\tau}A^*_s+\sum_{r=\tau+1}^{s\wedge(\tau+q)}\sum_{l=r-q}^{\tau}\Phi_{\tau l}U_{lr}\Phi^*_{sr}A^*_s$$

$$+\sum_{r=s-q}^{\tau}\Phi_{\tau r}W_{rs},$$

且由(4.1.3)与(4.1.9)有

$$
\begin{aligned}
&R(x_\tau,\ \hat{x}_{s|s-q-1})\\
&=R\left(x_\tau,\ \Phi_{s\tau}\hat{x}_{\tau|\tau-q}+\sum_{r=\tau+1}^{s-1}\Phi_{sr}\tilde{B}_{r-q}R(\varepsilon_{r-q})^{-}\varepsilon_{r-q}^{*}\right)\\
&=R(\hat{x}_{\tau|\tau-q})\Phi_{s\tau}^{*}+\sum_{r=\tau-q+1}^{s-q-1}B_{\tau r}R(\varepsilon_r)^{-}\tilde{B}_r^{*}\Phi_{s,r+q}^{*}.
\end{aligned}
$$

从而同样由(3.2.12)得到(4.1.23). ▌

为使用方便起见，我们列出以下两种特殊情形：

系 4.1.3　若(4.1.1)为白噪声系统，即 $q=0$，则其滤波、预测与平滑的递推公式如下：

$$
\begin{aligned}
\hat{x}_{t|t-1}&=\Phi_{t,t-1}\hat{x}_{t-1|t-1}+Eu_t,\\
P_{t|t-1}&=\Phi_{t,t-1}P_{t-1|t-1}\Phi_{t,t-1}^{*}+U_{tt},\\
\varepsilon_t&=z_t-A_t\hat{x}_{t|t-1}-Ev_t,\\
R(\varepsilon_t)&=A_tP_{t|t-1}A_t^{*}+A_tW_{tt}+W_{tt}^{*}A_t^{*}+V_{tt},\\
\tilde{B}_t&=P_{t|t-1}A_t^{*}+W_{tt},\\
\hat{x}_{t|t}&=\hat{x}_{t|t-1}+\tilde{B}_tR(\varepsilon_t)^{-}\varepsilon_t,\\
P_{t|t}&=P_{t|t-1}-\tilde{B}_tR(\varepsilon_t)^{-}\tilde{B}_t^{*}.
\end{aligned}
$$

当 $\tau>t$ 时，

$$
\hat{x}_{\tau|t}=\Phi_{\tau t}\hat{x}_{t|t}+\sum_{s=t+1}^{\tau}Eu_s,
$$

$$
P_{\tau|t}=\Phi_{\tau t}P_{t|t}\Phi_{\tau t}^{*}+\sum_{s=t+1}^{\tau}\Phi_{\tau s}U_{ss}\Phi_{\tau s}^{*}.
$$

当 $\tau<t$ 时，$\hat{x}_{\tau|t}$ 与 $P_{\tau|t}$ 仍按 (4.1.20) 与 (4.1.21) 每步再递推 $t-\tau$ 次得出，但此时 $B_{\tau s}$ 的计算简化为：

$$
B_{\tau s}=\left(P_{\tau|\tau}\Phi_{s\tau}^{*}-\sum_{r=\tau+1}^{s-1}B_{\tau r}R(\varepsilon_r)^{-}\tilde{B}_r^{*}\Phi_{sr}^{*}\right)A_s^{*},\ (s=\tau+1,\ \cdots,\ t).
$$

前面的 7 个公式就是熟知的 Kalman 滤波公式（通常考虑 $R(u_t,\ v_t)\triangleq W_{ts}\equiv 0$ 的情形）.

系 4.1.4　若系统(4.1.1)中的 $q=1$，且 $W_{ts}\equiv 0$，则其滤波公式为

$$
\begin{aligned}
\hat{x}_{t|t-2}&=\Phi_{t,t-1}\hat{x}_{t-1|t-2}+Eu_t,\\
P_{t|t-2}&=\Phi_{t,t-1}P_{t-1|t-2}\Phi_{t,t-1}^{*}+\Phi_{t,t-1}U_{t-1,t}+U_{t,t-1}\Phi_{t,t-1}^{*}+U_{tt},\\
\hat{x}_{t|t-1}&=\hat{x}_{t|t-2}+\tilde{B}_{t-1}R(\varepsilon_{t-1})^{-}\varepsilon_{t-1},
\end{aligned}
$$

$$P_{t|t-1}=P_{t|t-2}-\widetilde{B}_{t-1}R(\varepsilon_{t-1})^{-}\widetilde{B}_{t-1}^{*},$$

$$\varepsilon_t=z_t-A_t\hat{x}_{t|t-1}-V_{t,t-1}R(\varepsilon_{t-1})^{-}\varepsilon_{t-1}-Ev_t,$$

$$R(\varepsilon_t)=A_tP_{t|t-1}A_t^{*}+V_{tt}-V_{t,t-1}R(\varepsilon_{t-1})^{-}V_{t-1,t}$$
$$-A_t\widetilde{B}_{t-1}R(\varepsilon_{t-1})^{-}V_{t-1,t}-V_{t,t-1}R(\varepsilon_{t-1})^{-}\widetilde{B}_{t-1}^{*}A_t^{*},$$

$$B_{tt}=P_{t|t-1}A_t^{*}-\widetilde{B}_{t-1}R(\varepsilon_{t-1})^{-}V_{t-1,t},$$

$$\hat{x}_{t|t}=\hat{x}_{t|t-1}+B_{tt}R(\varepsilon_t)^{-}\varepsilon_t,$$

$$P_{t|t}=P_{t|t-1}-B_{tt}R(\varepsilon_t)^{-}B_{tt}^{*},$$

$$\widetilde{B}_t=\Phi_{t+1,t}B_{tt}+U_{t+1,t}A_t^{*}.$$

当 $\tau<t$ 时，$\hat{x}_{\tau|t}$、$P_{\tau|t}$ 仍按(4.1.20)与(4.1.21)每步再递推 $t-\tau$ 次得出，此时 $B_{\tau s}$ 的递推式为：对 $s=\tau+1,\cdots,t$,

$$B_{\tau s}=\left(P_{\tau|\tau-1}\Phi_{s\tau}^{*}+U_{\tau,\tau+1}\Phi_{s,\tau+1}^{*}-\sum_{r=\tau}^{s-1}B_{\tau r}R(\varepsilon_r)^{-}\widetilde{B}_r^{*}\Phi_{s,r+1}^{*}\right)A_s^{*}$$
$$-B_{\tau,s-1}R(\varepsilon_{s-1})^{-}V_{s-1,s}.$$

证 只须注意，在此情形由(4.1.5′)可算出 $J_{t,t-1}=A_t\widetilde{B}_{t-1}+V_{t,t-1}$. 将这一结果代入定理 4.1.1 和 4.1.2 中诸式并稍加整理即得. ∎

§4.2 含未知输入情形的滤波

假设与记号 4.2.1 在§1 中考虑系统(4.1.1)的滤波时，我们假定了系统不包含任何未知参数. 在本节中，我们进一步讨论 u_N, v_N 的均值函数含有未知线性回归系数向量 $\theta\in C^k$ 的情形，或称为有未知输入的情形. 对于 u_T、v_T 为白噪声这种特殊情形的讨论，可参见参考文献[15]、[16]、[20].

设对任意的 $t\in N$,

$$E_\theta u_t=G_t^{*}\theta,\ E_\theta v_t=H_t^{*}\theta,\ 且(G,\ H)\in\mathscr{R}\left(\begin{pmatrix}u\\v\end{pmatrix},\ t\right)^k.$$

这样，$\begin{pmatrix}u_N\\v_N\end{pmatrix}$ 对于任何 θ 都是局部正则的，因而由引理 3.1.3，z_N 也是局部正则序列. 为了与上一节的滤波记号有所区别，我们用

$\check{x}_{\tau|t}$ 记基于到 t 为止的量测 z_N, 对 x_τ 所作的 Gauss-Markov 估计(GM 滤波), $\Pi_{\tau|t} \triangleq R(x_\tau - \check{x}_{\tau|t})$. 又若 θ 有已知的先验一、二阶矩: $E_{ap}\theta = \mu_{ap}$, $R_{ap}(\theta) = \Gamma_{ap}$, 则用 $\check{x}_{\tau|t}$ 表其相应的线性 Bayes 估计(LB 滤波), $Q_{\tau|t} \triangleq E_{ap}E_\theta(x_\tau - \check{x}_{\tau|t})(x_\tau - \check{x}_{\tau|t})^*$(参见§2.3). 其中一种特别有用的情形是 $G_t = 0 (t > 1)$; $H_t \equiv 0$, 称为缺初值情形(这时只有状态初值 x_1 含未知参数 θ: $E_\theta x_1 = E_\theta u_1 = G_1\theta$).

根据上面的假设, 由(4.1.1)得到

$$E_\theta x_t = \sum_{s=1}^{t} \Phi_{ts} E_\theta u_s = \sum_{s=1}^{t} \Phi_{ts} G_s^* \theta,$$

$$E_\theta z_t = \left(A_t \sum_{s=1}^{t} \Phi_{ts} G_s^* + H_t^* \right) \theta.$$

再令

$$K_t \triangleq \sum_{s=1}^{t} G_s \Phi_{ts}^*, \quad F_t \triangleq K_t A_t^* + H_t, \tag{4.2.1}$$

则有

$$E_\theta x_t = K_t^* \theta, \quad E_\theta z_t = F_t^* \theta, \tag{4.2.2}$$

而诸协方差仍如(4.1.2).

定理 4.2.2 若系统(4.1.1)满足假设 4.2.1, 则对任意 $t \in N$, 有 $F \in \mathscr{R}_t^q$, 且 $z(F)_t$、$S(F)_t$ 满足如下递推方程:

$$\tilde{F}_t = (L_{t-1}^* \Phi_{t,t-1}^* + G_t) A_t^* + H_t - \sum_{s=t-q}^{t-1} \tilde{F}_s R(\varepsilon_s)^- J_{ts}^*, \tag{4.2.3}$$

$$L_t = \Phi_{t,t-1} L_{t-1} + G_t^* - \tilde{B}_{t-q} R(\varepsilon_{t-q})^- \tilde{F}_{t-q}^*, \tag{4.2.4}$$

$$z(F)_t = z(F)_{t-1} + \tilde{F}_t R(\varepsilon_t)^- \varepsilon_t, \tag{4.2.5}$$

$$S(F)_t = S(F)_{t-1} + \tilde{F}_t R(\varepsilon_t)^- \tilde{F}_t^*. \tag{4.2.6}$$

此外, 若存在 $t_0 \in N$, 使矩阵$(F_1, \cdots, F_{t_0})$为行线性无关, 则对任意的 $t \geqslant t_0$, $S(F)_t$ 为满秩阵, 且当 $t > t_0$ 时 $S(F)_t^{-1}$ 满足递推式

$$\begin{aligned} S(F)_t^{-1} = S(F)_{t-1}^{-1} - S(F)_{t-1}^{-1} \tilde{F}_t (R(\varepsilon_t) \\ + \tilde{F}_t^* S(F)_{t-1}^{-1} \tilde{F}_t)^- \tilde{F}_t^* S(F)_{t-1}^{-1}. \end{aligned} \tag{4.2.7}$$

其中 J_{ts}、ε_t、$R(\varepsilon_t)$、$\tilde{B}_{t-q}$ 由定理 4.1.1 算出, 但那里的 Eu_t、Ev_t 都用 0 代替.

证 $F_t \in \mathscr{R}_t^q$ 等价于 z_N 的局部正则性. 令

$$L_t \triangleq K_t^* - S(\Phi_t^{(q)} B, F)_{t-q},$$

其中根据 (3.2.2) 有 $(\Phi_t^{(q)}B)_s \triangleq \Phi_{t,s+q}R(x_{s+q},\ z_s)=R(x_t,\ z_s)$, $(s\in N_{t-q})$. 由(4.2.1)知 $K_t=K_{t-1}\Phi_{t,t-1}^*+G_t$. 再由 (3.2.12) 与 (3.2.13)即有

$$
\begin{aligned}
\tilde{F}_t &\triangleq R(z(F)_t,\ \varepsilon_t)\\
&=F_t-S(F,\ \Phi_{t-1}^{(q)}B)_{t-1-q}\Phi_{t,t-1}^*A_t^*-\sum_{s=t-q}^{t-1}\tilde{F}_sR(\varepsilon_s)^-J_{ts}^*\\
&=(K_{t-1}\Phi_{t,t-1}^*+G_t)A_t^*+H_t-S(F,\ \Phi_{t-1}^{(q)}B)_{t-1-q}\Phi_{t,t-1}^*A_t^*\\
&\quad-\sum_{s=t-q}^{t-1}\tilde{F}_sR(\varepsilon_s)^-J_{ts}^*\\
&=(L_{t-1}^*\Phi_{t,t-1}^*+G_t)A_t^*+H_t-\sum_{s=t-q}^{t-1}\tilde{F}_sR(\varepsilon_s)^-J_{ts}^*,\\
L_t&=\Phi_{t,t-1}K_{t-1}^*+G_t^*-\Phi_{t,t-1}S(\Phi_{t-1}^{(q)}B,\ F)_{t-1-q}\\
&\quad-\tilde{B}_{t-q}R(\varepsilon_{t-q})^-\tilde{F}_{t-q}^*\\
&=\Phi_{t,t-1}L_{t-1}+G_t^*-\tilde{B}_{t-q}R(\varepsilon_{t-q})^-\tilde{F}_{t-q}^*.
\end{aligned}
$$

于是证得(4.2.3)～(4.2.6).

关于定理的后一部分, 由定理2.2.2, (4.2.6) 及 $\tilde{F}_t=\tilde{F}_tR(\varepsilon_t)^-R(\varepsilon_t)$容易直接验证. ▌

定理 4.2.3 设系统(4.1.1)满足假设 4.2.1, 且存在 $t_0\in N$, 使矩阵 $(F_1,\ \cdots,\ F_{t_0})$ 为行线性无关. 又记 $\tau=t+\delta$, 其中 δ 为一固定整数, $L_{\tau t}\triangleq K_\tau^*-S(Z^{x_\tau},\ F)_t$. 则当 $t\geqslant t_0$ 时, GM 滤波 $\breve{x}_{\tau|t}$ 存在, 且 $\breve{x}_{\tau|t}$, $\Pi_{\tau|t}$ 由下式算出:

$$
\breve{x}_{\tau|t}=\hat{x}_{\tau|t}+L_{\tau t}S(F)_t^{-1}z(F)_t, \tag{4.2.8}
$$

$$
\Pi_{\tau|t}=P_{\tau|t}+L_{\tau t}S(F)_t^{-1}L_{\tau t}^*. \tag{4.2.9}
$$

其中 $L_{\tau t}$ 按下式递推:

① 当 $\delta\geqslant q$ 时,

$$
L_{\tau t}=\Phi_{\tau,\tau-1}L_{\tau-1,t-1}+G_\tau^*-\Phi_{\tau,t+q}\tilde{B}_tR(\varepsilon_t)^-\tilde{F}_t^*. \tag{4.2.10}
$$

② 当 $\delta<q$ 时, 在每一时刻 t 还须递推 $q-\delta$ 次: $L_{\tau,\tau-q}=L_\tau$,

$$
L_{\tau s}=L_{\tau,s-1}-R_{\tau s}R(\varepsilon_s)^-\tilde{F}_s^*,\quad s=\tau-q+1,\ \cdots,\ t. \tag{4.2.11}
$$

以上诸式中的 $\tilde{B}_t$、$R(\varepsilon_t)$; $\hat{x}_{\tau|t}$、$P_{\tau|t}$、$B_{\tau s}$; $\tilde{F}_t$、L_τ、$S(F)_t$、$z(F)_t$ 分别由定理 4.1.1, 4.1.2 与 4.2.2 递推算出, 但定理4.1.1与4.1.2

中的 Eu_t、Ev_t 都用 0 代替.

证 由系 2.3.3 与(4.1.11)即证得(4.2.8)与(4.2.9).

又当 $\delta\geqslant q$ 即 $\tau\geqslant t+q$ 时, 由(4.1.12)与(3.2.13),

$$L_{\tau t}=K_\tau^*-\Phi_{\tau,t+q}S(\Phi_{t+q}^{(q)}B,\ F)_t$$

$$=\Phi_{\tau,\tau-1}K_{\tau-1}+G_\tau^*-\Phi_{\tau,t+q}(\Phi_{t+q,t+q-1}S(\Phi_{t+q-1}^{(q)}B,\ F)_{t-1}$$

$$-\tilde{B}_tR(\varepsilon_t)^-\tilde{F}_t^*)$$

$$=\Phi_{\tau,\tau-1}L_{\tau-1,t-1}+G_\tau^*-\Phi_{\tau,t+q}\tilde{B}_tR(\varepsilon_t)^-\tilde{F}_t^*.$$

得到(4.2.10), 且知 $L_{t+q,t}=L_{t+q}$. 当 $\delta<q$ 即 $\tau<t+q$ 时, 由(4.1.20), 依次对 $s=\tau-q+1,\ \cdots,\ t$ 有

$$L_{\tau s}=K_\tau^*-R(\hat{x}_{\tau|s},\ z(F)_s)$$

$$=K_\tau^*-R(\hat{x}_{\tau|s-1}+B_{\tau s}R(\varepsilon_s)^-\varepsilon_s,\ z(F)_s)$$

$$=K_\tau^*-S(Z^{x_\tau},\ F)_{s-1}-B_{\tau s}R(\varepsilon_s)^-\tilde{F}_s^*$$

$$=L_{\tau,s-1}-B_{\tau s}R(\varepsilon_s)^-\tilde{F}_s^*.$$

这就是(4.2.11). ∎

定理 4.2.4 设系统(4.1.1)满足假设 4.2.1, 且 θ 有已知的先验一、二阶矩: $E_{ap}\theta=\mu_{ap}$, $R_{ap}(\theta)=\Gamma_{ap}$. 则 LB 滤波 $\check{x}_{\tau|t}$ 及其均方误差阵 $Q_{\tau|t}$ 由下式算出:

$$\check{x}_{\tau|t}=\hat{x}_{\tau|t}+L_{\tau t}[\Gamma_{ap}S(F)_t(S(F)_t+S(F)_t\Gamma_{ap}S(F)_t)^-(z(F)_t-S(F)_t\mu_{ap})+\mu_{ap}],\qquad(4.2.12)$$

$$Q_{\tau|t}=P_{\tau|t}+L_{\tau t}[\Gamma_{ap}-\Gamma_{ap}S(F)_t(S(F)_t+S(F)_t\Gamma_{ap}S(F)_t)^-S(F)_t\Gamma_{ap}]L_{\tau t}^*.\qquad(4.2.13)$$

特别是, 若 Γ_{ap} 为满秩阵, 则上两式可简化为:

$$\check{x}_{\tau|t}=\hat{x}_{\tau|t}+L_{\tau t}(\Gamma_{ap}^{-1}+S(F)_t)^{-1}(z(F)_t+\Gamma_{ap}^{-1}\mu_{ap}),\qquad(4.2.12')$$

$$Q_{\tau|t}=P_{\tau|t}+L_{\tau t}(\Gamma_{ap}^{-1}+S(F)_t)^{-1}L_{\tau t}^*.\qquad(4.2.13')$$

其中 $L_{\tau t}$ 的递推式仍如定理 4.2.3.

证 由定理 2.3.5 与 2.2.6 即得. ∎

§4.3 带ARMA量测噪声系统的滤波

在本节中，我们考虑不同于(4.1.1)的另一类非白噪声线性系统，即带时变 ARMA 量测噪声 v_N 的线性系统：

$$\begin{cases} x_t = \Phi_{t,t-1} x_{t-1} + u_t; \\ z_t = A_t x_t + v_t. \end{cases} \tag{4.3.1}$$

其中 u_N、v_N 的均值为0，且存在 $p \in N$，$\Psi_j(t) \in C^{m\times m}$，$(j=1, \cdots, p)$，使 $\bar{v}_t \triangleq v_t - \sum_{j=1}^{p} \Psi_j(t) v_{t-j}$（依通常约定，当 $s \leqslant 0$ 时，$v_s \triangleq 0$，$\bar{v}_s \triangleq 0$）对任意 $t \in N$ 满足

$$\begin{aligned} &R(u_t, u_s) = 0, \ R(u_t, v_s) = 0, \ s \in N_{t-q-1}, \\ &R(\bar{v}_t, \bar{v}_s) = 0, \ R(\bar{v}_t, u_s) = 0, \ s \in N_{t-p-q-1}. \end{aligned} \tag{4.3.2}$$

注意这两个条件表明了 u_N 为 q 步后不相关序列，而 v_N 为时变 ARMA$(p, p+q)$序列（参见定义 3.3.2）.

定理 4.3.1 在上述假设下，$\left(\bar{z}_t \triangleq z_t - \sum_{j=1}^{p} \Psi_j(t) z_{t-j}, \quad t \in N\right)$ 有 $p+q$ 步后可分解的协方差.

证 由上述假设，对任意的 $s \in N_{t-p-q-1}$，

$$\begin{aligned} R(\bar{z}_t, \bar{z}_s) &= R\left(A_t x_t - \sum_{j=1}^{p} \Psi_j(t) A_{t-j} x_{t-j} + \bar{v}_t, \ \bar{z}_s\right) \\ &= R\left(A_t\left(\Phi_{t,s+q} x_{s+q} + \sum_{r=s+q+1}^{t} \Phi_{tr} u_r \right)\right. \\ &\quad \left. - \sum_{j=1}^{p} \Psi_j(t) A_{t-j}\left(\Phi_{t-j,s+q} x_{s+q} + \sum_{r=s+q+1}^{t-j} \Phi_{t-j,r} u_r \right), \bar{z}_s\right) \\ &= \left(A_t \Phi_{t,s+q} - \sum_{j=1}^{p} \Psi_j(t) A_{t-j} \Phi_{t-j,s+q}\right) R(x_{s+q}, \ \bar{z}_s). \end{aligned}$$

对 $r \in N_t$，记

$$A_{tr} \triangleq A_t \Phi_{tr} - \sum_{j=1}^{(t-r)\wedge p} \Psi_j(t) A_{t-j} \Phi_{t-j,r}. \tag{4.3.3}$$

则从上式即得

$$R(\bar{z}_t, \ \bar{z}_s) = A_{t,t-p} \Phi_{t-p,s+q} R(x_{s+q}, \ \bar{z}_s). \tag{4.3.4}$$

对比定义 3.2.1 中的(3.2.3)式，若暂记 $\overline{A}_t \triangleq A_{t,t-p}$，$\overline{\Phi}_{t,t-1} \triangleq \Phi_{t-p,t-p-1}$，$\overline{B}_s \triangleq R(x_{s+q}, \bar{z}_s)$，则(4.3.4)又可表成

$$R(\bar{z}_t, \bar{z}_s) = \overline{A}_t \overline{\Phi}_{t,s+p+q} \overline{B}_s, \quad s \in N_{t-p-q-1}.$$

又对任意的 $s \in N_{t-p-q}$，

$$\begin{aligned}(\overline{\Phi}_t^{(p+q)} \overline{B})_s &\triangleq \overline{\Phi}_{t,s+p+q} \overline{B}_s = \Phi_{t-p,s+q} R(x_{s+q}, \bar{z}_s) \\ &= R\left(\Phi_{t-p,s+q} x_{s+q} + \sum_{r=s+q+1}^{t-p} \Phi_{t-p,r} u_r, \ \bar{z}_s\right) \\ &= R(x_{t-p}, \bar{z}_s). \end{aligned} \tag{4.3.5}$$

所以 $\overline{\Phi}_t^{(p+q)} \overline{B} \in \mathscr{R}(\bar{z}, t-p-q)^*$. 故依定义 3.2.1，$\bar{z}_N$ 对 $\overline{\Phi} \triangleq \{\overline{\Phi}_{t,t-1}, t \in N\}$ 有 $p+q$ 步后可分解的协方差. ▮

由定理 3.3.1，对任意 $t \in N$，$\mathscr{H}(\bar{z}, t) = \mathscr{H}(z, t) \triangleq \mathscr{H}_t$，$\bar{z}_N$ 的新息序列与 z_N 的新息序列相同，且 $\mathscr{R}(\bar{z}, t)$ 与 $\mathscr{R}(z, t) \triangleq \mathscr{R}_t$ 存在同构对应. 因此根据上面证明的 $\bar{z}_N$ 的协方差可分解性，用定理 3.2.3 即可解决有关 z_N 的所有线性统计递推计算问题，特别是系统(4.3.1)的滤波问题. 为简单起见，下面我们只考虑 $q=0$ 的情形.

定理 4.3.2 对于满足(4.3.2)的线性系统(4.3.1)，设其中的 $q=0$. 记

$$\begin{aligned} &U_{ts} \triangleq R(u_t), \ \overline{V}_{ts} \triangleq R(\bar{v}_t, \bar{v}_s), \\ &\overline{W}_{st} \triangleq R(u_s, \bar{v}_t), \ (t-p \leqslant s \leqslant t). \end{aligned} \tag{4.3.6}$$

并简记 $\hat{x}_t \triangleq \hat{x}_{t|t}$，$P_t \triangleq P_{t|t}$. 则其新息与滤波有如下递推式：

$$\begin{aligned} J_{ts} = {} & A_{ts} \overline{B}_s + (A_{t,s-p-1} - A_{ts} \Phi_{s,s-p-1}) P_{s-p-1} A^*_{s,s-p-1} \\ & + \sum_{r=t-p}^{s} [\overline{W}^*_{rt} A^*_{sr} + (A_{tr} - A_{ts} \Phi_{sr})(U_{rr} A^*_{sr} + \overline{W}_{rs})] \\ & + \overline{V}_{ts} - \sum_{r=t-p}^{s-1} (J_{tr} - A_{ts} \Phi_{sr} \overline{B}_r) R(\varepsilon_r)^- J^*_{sr}, \\ & (s = t-p, \cdots, t-1), \end{aligned} \tag{4.3.7}$$

$$\varepsilon_t = z_t - \sum_{j=1}^{p} \Psi_j(t) z_{t-j} - A_{t,t-p-1} \hat{x}_{t-p-1} - \sum_{s=t-p}^{t-1} J_{ts} R(\varepsilon_s)^- \varepsilon_s, \tag{4.3.8}$$

$$
\begin{aligned}
R(\varepsilon_t)=&A_{t,t-p-1}P_{t-p-1}A^*_{t,t-p-1}\\
&+\sum_{s=t-p}^{t}(A_{ts}U_{ss}A^*_{ts}+A_{ts}\overline{W}_{st}+\overline{W}^*_{st}A^*_{ts})\\
&+\overline{V}_{tt}-\sum_{s=t-p}^{t-1}J_{ts}R(\varepsilon_s)^-J^*_{ts},
\end{aligned}
\tag{4.3.9}
$$

$$
\begin{aligned}
\widetilde{B}_t=&\Phi_{t,t-p-1}P_{t-p-1}A^*_{t,t-p-1}+\sum_{s=t-p}^{t}\Phi_{ts}(U_{ss}A^*_{ts}+\overline{W}_{st})\\
&-\sum_{s=t-p}^{t-1}\Phi_{ts}\widetilde{B}_sR(\varepsilon_s)^-J^*_{ts},
\end{aligned}
\tag{4.3.10}
$$

$$\hat{x}_t=\Phi_{t,t-1}\hat{x}_{t-1}+\widetilde{B}_tR(\varepsilon_t)^-\varepsilon_t, \tag{4.3.11}$$

$$P_t=\Phi_{t,t-1}P_{t-1}\Phi^*_{t,t-1}+U_{tt}-\widetilde{B}_tR(\varepsilon_t)^-\widetilde{B}^*_t. \tag{4.3.12}$$

此外，对 $\tau>t$ 还有下式成立：

$$\hat{x}_{\tau|t}=\Phi_{\tau,\tau-1}\hat{x}_{\tau-1|t}=\Phi_{\tau t}\hat{x}_t, \tag{4.3.13}$$

$$P_{\tau|t}=\Phi_{\tau,\tau-1}P_{\tau-1|t}\Phi^*_{\tau,\tau-1}+U_{tt}. \tag{4.3.14}$$

证　因为考虑 $q=0$ 的情形，由定理 4.3.1, $\bar{z}_N$ 对 $\overline{\Phi}$ 有 p 步后可分解的协方差. 又由(4.3.5)，对任意 $s\in N_{t-p}$ 有

$$(\overline{\Phi}_t^{(p)}\overline{B})_s=R(x_{t-p},\ \bar{z}_s),$$

所以有 $\bar{z}(\overline{\Phi}_t^{(p)}\overline{B})_{t-p}=\hat{x}_{t-p}$ 或者 $\bar{z}(\overline{\Phi}_{t+p}^{(p)}\overline{B})_t=\hat{x}_t$. 当 $\tau>t$ 时，由引理 3.2.2 即知，$\hat{x}_{\tau|t}=\bar{z}(\overline{\Phi}_{\tau+p}^{(p)}\overline{B})_t=\overline{\Phi}_{\tau+p,t+p}\bar{z}(\overline{\Phi}_{t+p}^{(p)}\overline{B})_t=\Phi_{\tau t}\hat{x}_t=\Phi_{\tau,\tau-1}\hat{x}_{\tau-1|t}$, $R(\hat{x}_{\tau|t})=\Phi_{\tau,\tau-1}R(\hat{x}_{\tau-1|t})\Phi^*_{\tau,\tau-1}$. 再注意递推式 $R(x_\tau)=R(\Phi_{\tau,\tau-1}x_{\tau-1}+u_\tau)=\Phi_{\tau,\tau-1}R(x_{\tau-1})\Phi^*_{\tau,\tau-1}+U_{\tau\tau}$ 以及 $P_{\tau|t}=R(x_\tau)-R(\hat{x}_{\tau|t})$，即证得(4.3.13)与(4.3.14).

又依序可以算出，对任意的 $t-p\leqslant s\leqslant t$，有

$$
\begin{aligned}
R(u_s,\ \bar{z}_t)&=R\Big(u_s,\ \ A_tx_t-\sum_{j=1}^{p}\Psi_j(t)A_{t-j}x_{t-j}+\bar{v}_t\Big)\\
&=R\Big(u_s,\ \Big(A_t\Phi_{ts}-\sum_{j=1}^{t-s}\Psi_j(t)A_{t-j}\Phi_{t-j,s}\Big)u_s\Big)+\overline{W}_{st}\\
&=U_{ss}A^*_{ts}+\overline{W}_{st},
\end{aligned}
$$

$$
\begin{aligned}
R(x_s,\ \bar{z}_t)&=R\Big(\Phi_{s,t-p-1}x_{t-p-1}+\sum_{r=t-p}^{s}\Phi_{sr}u_r,\ \bar{z}_t\Big)\\
&=\Phi_{s,t-p-1}R\Big(x_{t-p-1},\ A_tx_t-\sum_{j=1}^{p}\Psi_j(t)A_{t-j}x_{t-j}+\bar{v}_t\Big)
\end{aligned}
$$

$$
\begin{aligned}
&+\sum_{r=t-p}^{t}\Phi_{sr}R(u_r,\ \bar{z}_t)\\
&=\Phi_{s,t-p-1}R(x_{t-p-1})A^*_{t,t-p-1}+\sum_{r=t-p}^{s}\Phi_{sr}(U_{rr}A^*_{tr}+\overline{W}_{rt}),
\end{aligned}
$$

$$
\begin{aligned}
R(\bar{z}_s,\ \bar{z}_t)&=A_{ts}R(x_s,\ \bar{z}_s)-\sum_{j=t-s+1}^{p}\Psi_j(t)A_{t-j}R(x_{t-j},\ \bar{z}_s)\\
&\quad+R(\bar{v}_t,\ \bar{z}_s)\\
&=A_{ts}\Big[\Phi_{s,s-p-1}R(x_{s-p-1})A^*_{s,s-p-1}\\
&\quad+\sum_{r=s-p}^{s}\Phi_{sr}(U_{rr}A^*_{sr}+\overline{W}_{rs})\Big]\\
&\quad-\sum_{j=t-s+1}^{p}\Psi_j(t)A_{t-j}\Big[\Phi_{t-j,s-p-1}R(x_{s-p-1})A^*_{s,s-p-1}\\
&\quad+\sum_{r=s-p}^{t-j}\Phi_{t-j,r}(U_{rr}A^*_{sr}+\overline{W}_{rs})\Big]\\
&\quad+R\Big(\bar{v}_t,\ A_sx_s-\sum_{j=1}^{p}\Psi_j(s)A_{s-j}x_{s-j}+\bar{v}_s\Big)\\
&=A_{t,s-p-1}R(x_{s-p-1})A^*_{s,s-p-1}\\
&\quad+\sum_{r=s-p}^{s}A_{ts}\Phi_{sr}(U_{rr}A^*_{sr}+\overline{W}_{rs})\\
&\quad-\sum_{r=s-p}^{s-1}\sum_{j=t-s+1}^{(t-r)\wedge p}\Psi_j(t)A_{t-j}\Phi_{t-j,r}(U_{rr}A^*_{sr}+\overline{W}_{rs})\\
&\quad+\sum_{r=t-p}^{s}\overline{W}^*_{rt}A^*_{sr}+\overline{V}_{ts}\\
&=A_{t,s-p-1}R(x_{s-p-1})A^*_{s,s-p-1}\\
&\quad+\sum_{r=s-p}^{s}A_{tr}(U_{rr}A^*_{sr}+\overline{W}_{rs})+\sum_{r=t-p}^{s}\overline{W}^*_{rt}A^*_{sr}+\overline{V}_{ts}.
\end{aligned}
$$

将以上结果代入定理 3.2.3 中(3.2.7)～(3.2.11)诸式，并注意那里的 $q=p$，$y_t=\tilde{z}(\overline{\Phi}^{(p)}_t\overline{B})_{t-p}=\hat{x}_{t-p}$，$\overline{\Phi}_{ts}=\Phi_{t-p,s-p}$，$\overline{A}_t=A_{t,t-p}$，$\overline{R}_t=R(x_t,\ \bar{z}_t)$，即证得(4.3.8)～(4.3.12)。至于(4.3.7)，由(3.2.6)，对 $s=t-p,\ \cdots,\ t-1$，

$$
\begin{aligned}
J_{ts}&=R(\bar{z}_t,\ \bar{z}_s)-A_{t,s-p-1}R(\hat{x}_{s-p-1})A^*_{s,s-p-1}\\
&\quad-\sum_{r=s-p}^{t-p-1}A_{tr}\tilde{B}_rR(\varepsilon_r)^-J^*_{sr}-\sum_{r=t-p}^{s-1}J_{tr}R(\varepsilon_r)^-J^*_{sr}
\end{aligned}
$$

$$-A_{t,s-p-1}P_{s-p-1}A^*_{s,s-p-1}+\sum_{r=s-p}^{s}A_{tr}(U_{rr}A^*_{sr}+\overline{W}_{rs})$$
$$+\sum_{r=t-p}^{s}\overline{W}^*_{rt}A^*_{sr}+\overline{V}_{ts}-\sum_{r=s-p}^{t-p-1}A_{tr}\widetilde{B}_rR(\varepsilon_r)^-J^*_{sr}$$
$$-\sum_{r=t-p}^{s-1}J_{tr}R(\varepsilon_r)^-J^*_{sr}.$$

注意当 $r\leqslant t-p$ 时, $A_{tr}=A_{t,t-p}\Phi_{t-p,r}$, 再由(4.3.10), 即易将其化为更便于递推的形式(4.3.7). ▮

为使用方便起见, 我们列出如下 $q=0$, $p=1$ 的特殊情形.

系 4.3.3　在线性系统(4.3.1)中, 设

$$R(u_t,\ u_s)=0,\ R(u_t,\ \overline{v}_s)=0,\ R(\overline{v}_t,\ \overline{v}_s)=0,\ (t\neq s).$$

其中 $\overline{v}_t=v_t-\Psi(t)v_{t-1}$. 则其新息与滤波有如下递推公式:

$$\varepsilon_t=z_t-\Psi(t)z_{t-1}-(A_t\Phi_{t,t-1}-\Psi(t)A_{t-1})\hat{x}_{t-1},$$
$$R(\varepsilon_t)=(A_t\Phi_{t,t-1}-\Psi(t)A_{t-1})P_{t-1}(A_t\Phi_{t,t-1}-\Psi(t)A_{t-1})^*$$
$$+A_tU_{tt}A^*_t+A_t\overline{W}_{tt}+\overline{W}^*_{tt}A^*_t+\overline{V}_{tt},$$
$$\widetilde{B}_t=\Phi_{t,t-1}P_{t-1}(A_t\Phi_{t,t-1}-\Psi(t)A_{t-1})^*+U_{tt}A^*_t+\overline{W}_{tt},$$
$$\hat{x}_t=\Phi_{t,t-1}\hat{x}_{t-1}+\widetilde{B}_tR(\varepsilon_t)^-\varepsilon_t,$$
$$P_t=\Phi_{t,t-1}P_{t-1}\Phi^*_{t,t-1}+U_{tt}-\widetilde{B}_tR(\varepsilon_t)^-\widetilde{B}^*_t.$$

证　注意对此特殊情形, $A_{t,t-1}=A_t\Phi_{t,t-1}-\Psi(t)A_{t-1}$, 且易由(4.3.7)得 $J_{t,t-1}=A_{t,t-1}\widetilde{B}_{t-1}$. 再利用(4.3.11)与(4.3.12)即易求出各式. ▮

最后, 我们考虑系统(4.3.1)含未知输入情形的 GM 与 LB 滤波.

定理 4.3.4　若系统(4.3.1)满足假设 4.2.1, 且 $q=0$. 则对任意 $t\in N$ 有 $F\in\mathscr{H}_t$, 且 $z(F)_t$、$S(F)_t$ 满足如下递推方程:

$$\widetilde{F}_t=L^*_{t-p-1}A^*_{t,t-p-1}+\sum_{s=t-p}^{t}G_sA^*_{ts}+H_t$$
$$-\sum_{j=1}^{p}H_{t-j}\Psi^*_j(t)-\sum_{s=t-p}^{t-1}\widetilde{F}_sR(\varepsilon_s)^-J^*_{ts},\qquad(4.3.15)$$
$$L_t=\Phi_{t,t-1}L_{t-1}+G^*_t-\widetilde{B}_tR(\varepsilon_t)^-\widetilde{F}^*_t,\qquad(4.3.16)$$
$$z(F)_t=z(F)_{t-1}+\widetilde{F}_tR(\varepsilon_t)^-\varepsilon_t,\qquad(4.3.17)$$

$$S(F)_t=S(F)_{t-1}+\widetilde{F}_tR(\varepsilon_t)^-\widetilde{F}_t^*. \tag{4.3.18}$$

此时 x_t 的 LB 滤波 $\check{x}_t$ 及其均方误差阵 Q_t 由下式算出：

$$\begin{aligned}\check{x}_t=\hat{x}_t+L_t[\Gamma_{ap}S(F)_t(S(F)_t\\+S(F)_t\Gamma_{ap}S(F)_t)^-(z(F)_t-S(F)_t\mu_{ap})+\mu_{ap}],\end{aligned} \tag{4.3.19}$$

$$\begin{aligned}Q_t=P_t+L_t[\Gamma_{ap}S(F)_t(S(F)_t\\+S(F)_t\Gamma_{ap}S(F)_t)^-S(F)_t\Gamma_{ap}]L_t^*.\end{aligned} \tag{4.3.20}$$

若 Γ_{ap} 为满秩阵，则以上两式更可简化为

$$\check{x}_t=\hat{x}_t+L_t(\Gamma_{ap}^{-1}+S(F)_t)^{-1}(z(F)_t+\Gamma_{ap}^{-1}\mu_{ap}), \tag{4.3.19'}$$

$$Q_t=P_t+L_t(\Gamma_{ap}^{-1}+S(F)_t)^{-1}L_t^*. \tag{4.3.20'}$$

若进一步设存在 $t_0\in N$ 使矩阵 $(F_1,\cdots,F_{t_0})$ 为行线性无关，则对任意 $t\geqslant t_0$，$S(F)_t$ 为满秩阵，且当 $t>t_0$ 时 $S(F)_t^{-1}$ 满足递推方程

$$\begin{aligned}S(F)_t^{-1}=S(F)_{t-1}^{-1}-S(F)_{t-1}^{-1}\widetilde{F}_t(R(\varepsilon_t)\\+\widetilde{F}_t^*S(F)_{t-1}\widetilde{F}_t)^-\widetilde{F}_t^*S(F)_{t-1}^{-1}.\end{aligned} \tag{4.3.21}$$

此时 $x_t(t\geqslant t_0)$ 的 GM 滤波 $\check{\check{x}}_t$ 及其误差方差阵 Π_t 由下式算出：

$$\check{\check{x}}_t=\hat{x}_t+L_tS(F)_t^{-1}z(F)_t, \tag{4.3.22}$$

$$\Pi_t=P_t+L_tS(F)_t^{-1}L_t^*. \tag{4.3.23}$$

以上诸式中的 J_{ts}、ε_t、$R(\varepsilon_t)$、$\widetilde{B}_t$、$\hat{x}_t$、P_t 都按定理 4.3.2 递推计算.

证　记

$$\widetilde{F}_t\triangleq F_t-\sum_{j=1}^{p}F_{t-j}\Psi_j^*(t),$$

由(4.2.2)，$E_\theta z_t=F_t^*\theta$，故 $E_\theta\bar{z}_t=\widetilde{F}_t^*\theta$. 由定理 3.3.1，$F\in\mathscr{R}_t^*$ 当且仅当 $\widetilde{F}\in\mathscr{R}(\bar{z},t)^k$，且 $z(F)_t=\bar{z}(\widetilde{F})_t$，$S(F)_t=S_{\bar{z}}(\widetilde{F})_t$. 又令 $L_t\triangleq K_t^*-S_{\bar{z}}(\overline{\Phi}_{t+p}^{(p)}\overline{B},\ \widetilde{F})_t$. 由 (4.3.5)，$L_t=K_t^*-R(\hat{x}_t,\ \bar{z}(\widetilde{F})_t)=K_t^*-R(\hat{x}_t,\ z(F)_t)=K_t^*-S(Z^{x_t},\ F)_t$.

由(4.2.1)，

$$\begin{aligned}\widetilde{F}_t=K_tA_t^*+H_t-\sum_{j=1}^{p}(K_{t-j}A_{t-j}^*+H_{t-j})\Psi_j^*(t)\\-\Big(K_{t-p-1}\Phi_{t,t-p-1}^*+\sum_{s=t-p}^{t}G_s\Phi_{ts}^*\Big)A_t^*\end{aligned}$$

$$-\sum_{j=1}^{p}\left(K_{t-p-1}\Phi^*_{t-j,t-p-1}+\sum_{s=t-p}^{t-j}G_s\Phi^*_{t-j,s}\right)A^*_{t-j}\Psi^*_j(t)+\overline{H}_t$$

$$=K_{t-p-1}A^*_{t,t-p-1}+\sum_{s=t-p}^{t}G_s\Phi^*_{ts}A^*_t$$

$$-\sum_{s=t-p}^{t-1}G_s\sum_{j=1}^{t-s}\Phi^*_{t-j,s}A^*_{t-j}\Psi^*_j(t)+\overline{H}_t$$

$$=K_{t-p-1}A^*_{t,t-p-1}+\sum_{s=t-p}^{t}G_sA^*_{ts}+\overline{H}_t.$$

再注意 $\bar{z}_N$ 对 $\overline{\Phi}$ 有 p 步后可分解的协方差，所以由 (3.2.12) 与 (3.2.13)得

$$\widetilde{F}_t \triangleq R(z(F)_t,\ \varepsilon_t)$$

$$=R(\bar{z}(\overline{F})_t,\ \varepsilon_t)$$

$$=\overline{F}_t-S_2(\overline{F},\ \overline{\Phi}^{(p)}_{t-1}\overline{B})_{t-p-1}A^*_{t,t-p-1}-\sum_{s=t-p}^{t-1}\widetilde{F}_sR(\varepsilon_s)^-J^*_{ts}$$

$$=L^*_{t-p-1}A^*_{t,t-p-1}+\sum_{s=t-p}^{t}G_sA^*_{ts}+\overline{H}_t-\sum_{s=t-p}^{t-1}\widetilde{F}_sR(\varepsilon_s)^-J^*_{ts},$$

$$L_t=\Phi_{t,t-1}K^*_{t-1}+G^*_t-\Phi_{t,t-1}S_2(\overline{\Phi}^{(p)}_{t+p-1}\overline{B},\ \overline{F})_{t-1}$$

$$-\widetilde{B}_tR(\varepsilon_t)^-\widetilde{F}^*_t$$

$$=\Phi_{t,t-1}L_{t-1}+G^*_t-\widetilde{B}_tR(\varepsilon_t)^-\widetilde{F}^*_t.$$

这就证明了(4.3.15)～(4.3.18). 至于定理其他部分的证明，则完全同于定理 4.2.3 及 4.2.4. ▌

第 5 章

随机积分与正交随机测度

随机积分是随机分析、也是连续时间随机过程统计的重要工具，很多有关著作中都有讨论.本章的特点是用再生核表示来定义与研究随机积分．为了更拓广些，我们将不限于实轴，而在一般可测空间上来讨论两种随机积分．为此，在阅读本章有些论述时，需要有测度论的基础知识.不具备这一条件的读者可以跳过本章，大致上不会影响对以后诸章基本内容的阅读.

§5.1 二阶可测过程与阵测度

定义 5.1.1 设$(T, \mathscr{T})$为可测空间，其中$\mathscr{T}$是T的子集组成的σ域，z_T为m维正则过程．若对任意的$c\in C^m$，$t\mapsto c^* z_t$是$(T, \mathscr{T})$映入$(\mathscr{H}(z_T), \mathscr{B}(\mathscr{H}(z_T)))$的可测映射，其中$\mathscr{B}(\mathscr{H}(z_T))$表示由 Hilbert 空间$(\mathscr{H}(z_T), B)$中的全体开集生成的$\sigma$域，则称$z_T$为$(T, \mathscr{T})$上的二阶可测过程.

由可测映射的基本知识可知，上述定义即相当于：对任意$x\in \mathscr{H}(z_T)$和$\alpha>0$有

$$\{t; B(c^* z_t - x)<\alpha\}\in \mathscr{T}.$$

定理 5.1.2 z_T为$(T, \mathscr{T})$上的二阶可测过程，当且仅当其协方差阵函数$B(z_s, z_t)(s、t\in T)$为$(T\times T, \mathscr{T}\times\mathscr{T})$上的可测阵函数．且此时$\mathscr{R}(z_T)^k$中的任一元必为$(T, \mathscr{T})$上的可测$k\times m$阵函数.

证 必要性：设z_T为二阶可测过程，则$T\times T\to C^{m\times m}$的映射$(s, t)\mapsto B(z_s, z_t)$是如下两可测映射的复合：

$$(s,\ t)\mapsto(z_s,\ z_t)(T\times T\to\mathscr{H}(z_T)^m\times\mathscr{H}(z_T)^m),$$

$$(x,\ y)\mapsto R(x,\ y)(\mathscr{H}(z_T)^m\times\mathscr{H}(z_T)^m\to C^{m\times m}).$$

其中后一映射为连续映射，因而是可测的.

充分性：设$(R(z_s,\ z_t),\ s、t\in T)$为二元可测，则由熟知的单调类定理，易证$(R(z_t),\ t\in T)$为一元可测（首先对$S_1、S_2\in\mathscr{T}$，考虑形如$(f(s,\ t)\triangleq 1_{S_1\times S_2}(s,\ t),\ s、t\in T)$的二元可测函数，此时$f(t,\ t)=1_{S_1\cap S_2}(t)$）. 又由二元可测函数的截口及可测函数列的极限的可测性易知，对任意$x\in\mathscr{H}(z_T)^k$，阵函数$(R(x,\ z_t),\ t\in T)$为可测，即$R(z_T)^k$中任一元为可测. 最后，对任意$c\in C^m$和$x\in\mathscr{H}(z_T)$有

$$R(c^*z_t-x)=c^*R(z_t)c-c^*R(z_t,\ x)-R(x,\ z_t)c+R(x).$$

右边诸项都是t的可测函数，所以对任意$\alpha>0$有$\{t:R(c^*z_t-x)<\alpha\}\in\mathscr{T}$. ▌

定义5.1.3 设$(T,\ \mathscr{T})$为一可测空间，若对任意的$S\in\mathscr{T}$，有一m阶非负定Hermite阵$\mu(S)$与之对应，即$\mu(S)^*=\mu(S)$，$\mu(S)\geqslant 0$，且对任意有限或可数个不交的$S_j\in\mathscr{T}$，有下式成立：

$$\mu(\sum_j S_j)=\sum_j\mu(S_j)$$

（其中左边的Σ表示集的不交并），则称μ为$(T,\ \mathscr{T})$上的m阶阵测度，$(T,\ \mathscr{T},\ \mu)$称为m阶阵测度空间.

若容许某些$\mu(S)$的元取正无穷值，但存在$(T_n)\subset\mathscr{T}$，$T_n\uparrow T$（即(T_n)为上升子集列，且$\bigcup_n T_n=T$），使每一$\mu(T_n)$的元均为有限，则称μ为$(T,\ \mathscr{T})$上的m阶σ有限阵测度.

引理5.1.4 设μ为$(T,\ \mathscr{T})$上的m阶（σ有限）阵测度，则有

i) μ的任一元为$(T,\ \mathscr{T})$上的有限（σ有限）变号测度；

ii) 任意$c\in C^m$，$c^*\mu c$为$(T,\ \mathscr{T})$上的有限（σ有限）测度；

iii) $\mathrm{tr}\mu$为$(T,\ \mathscr{T})$上的有限（σ有限）测度，称为μ的迹测度；

iv) 对任意$S\in\mathscr{T}$，$\mu(S)\leqslant[\mathrm{tr}\mu(S)]I_m$；

v) 设$S\in\mathscr{T}$，则$\mu(S)=0$当且仅当$\mathrm{tr}\mu(S)=0$，从而存在

Hermite 阵值的 Radon-Nikodym 导数 $\frac{d\mu}{d\operatorname{tr}\mu}$, 且 $0\leqslant\frac{d\mu}{d\operatorname{tr}\mu}\leqslant I_m$, a. e. $\operatorname{tr}\mu$ (即除去一 $\operatorname{tr}\mu$ 零测集外此不等式处处成立，或称对 $\operatorname{tr}\mu$ 几乎处处成立).

证　以上事实均为显然. ▌

定义 5.1.5　设 $(T, \mathscr{T}, \mu)$ 为 m 阶 σ 有限阵测度空间，F 和 G 分别为 $(T, \mathscr{T})$ 上的可测 $k\times m$ 和 $l\times m$ 阵函数. 根据引理 5.1.4, 我们可按矩阵乘法规则定义积分

$$\int_T F_t d\mu \quad 及 \quad \int_T F_t d\mu G_t^*$$

分别为 $k\times m$ 和 $k\times l$ 阵，如果其中每一元的每一项普通积分都存在且取有限值. 又易知有下式成立:

$$\int_T F_t d\mu = \int_T F_t \left(\frac{d\mu}{d\operatorname{tr}\mu}\right)_t d\operatorname{tr}\mu,$$

$$\int_T F_t d\mu G_t^* = \int_T F_t \left(\frac{d\mu}{d\operatorname{tr}\mu}\right)_t G_t^* d\operatorname{tr}\mu.$$

特别，积分 $\int_T F_t d\mu F_t^*$ 为一 k 阶非负定 Hermite 阵，若它取有限值 $\left(等价于 \operatorname{tr}\left(\int_T F_t d\mu F_t^*\right)<\infty\right)$, 则称 F 对 μ 平方可积，记作 $F\in L_2(T, \mathscr{T}, \mu)^k$, 或简记作 $F\in L_2(\mu)^k$. 若 $\int_T F_t d\mu F_t^* = 0$ (等价于 $\operatorname{tr}\left(\int_T F_t d\mu F_t^*\right)=0$), 则记作 $F=0$, a.e. μ. $F=G$, a. e. μ 意味着 $F-G=0$, a. e. μ. 若 $(F_n)\subset L_2(\mu)^k$, 且当 $n\to\infty$ 时 $\int_T (F_{nt}-F_t)d\mu(F_{nt}-F_t)^*\to 0$, 则记作 $F_n \xrightarrow{L_2(\mu)} F$.

又若存在 $(T_n)\subset\mathscr{T}$, $T_n\uparrow T$, 使 $\operatorname{tr}\left(\int_{T_n} F_t d\mu F_t^*\right)<\infty$, 则称 F 对 μ σ 平方可积，记作 $F\in\sigma L_2(\mu)^k$.

下面的引理，可视为普通积分的 Schwarz 不等式与 Cauchy 不等式的推广.

引理 5.1.6　设 $(T, \mathscr{T}, \mu)$ 为 m 阶 σ 有限阵测度空间，$F\in$

$L_2(\mu)^k$, $G\in L_2(\mu)^l$. 则有以下不等式成立:

$$\left(\int_T F_t d\mu G_t^*\right)\left(\int_T G_t d\mu G_t^*\right)^-\left(\int_T G_t d\mu F_t^*\right)$$
$$\leqslant\int_T F_t d\mu F_t^*, \tag{5.1.1}$$

$$\left\|\int_T F_t d\mu G_t^*\right\|^2\leqslant \operatorname{tr}\left(\int_T F_t d\mu F_t^*\right)\operatorname{tr}\left(\int_T G_t d\mu G_t^*\right). \tag{5.1.1'}$$

(5.1.1)取等号的必要充分条件是: 存在 $C\in C^{k\times l}$ 使 $F=CG$, a. e. μ. 若 $k=l$, 则还有以下不等式成立:

$$\left(\operatorname{tr}\left[\int_T (F_t+G_t)d\mu(F_t+G_t)^*\right]\right)^{\frac{1}{2}}$$
$$\leqslant\left(\operatorname{tr}\left(\int_T F_t d\mu F_t^*\right)\right)^{\frac{1}{2}}+\left(\operatorname{tr}\left(\int_T G_t d\mu G_t^*\right)\right)^{\frac{1}{2}}. \tag{5.1.2}$$

证 (5.1.1)式可由下式得出:

$$0\leqslant\int_T\left[F_t-\left(\int_T F_s d\mu G_s^*\right)\left(\int_T G_s d\mu G_s^*\right)^- G_t\right]d\mu$$
$$\times\left[F_t-\left(\int_T F_s d\mu G_s^*\right)\left(\int_T G_s d\mu G_s^*\right)^- G_t\right]^*$$
$$=\int_T F_t d\mu F_t^*$$
$$-\left(\int_T F_t d\mu G_t^*\right)\left(\int_T G_t d\mu G_t^*\right)^-\left(\int_T G_t d\mu F_t^*\right).$$

若上式$=0$, 则有 $F=\left(\int_T F_s d\mu G_s^*\right)\left(\int_T G_s d\mu G_s^*\right)^- G$, a.e.$\mu$; 反之, 若存在 $C\in C^{k\times l}$ 使 $F=CG$, a. e. μ, 则(5.1.1)式显然成为等式.

注意: 对 $g\in L_2(\mu)$ 有

$$\left(\int_T g_t d\mu g_t^*\right)^-=\begin{cases}0, & g=0,\ \text{a. e. } \mu;\\ 1\Big/\left(\int_T g_t d\mu g_t^*\right), & \text{否}.\end{cases}$$

故由(5.1.1)推出

$$\left(\int_T F_t d\mu g_t^*\right)\left(\int_T g_t d\mu F_t^*\right)\leqslant\left(\int_T g_t d\mu g_t^*\right)\left(\int_T F_t d\mu F_t^*\right),$$

两边求迹即得

$$\left\|\int_T F_t d\mu g_t^*\right\|^2 \leqslant \left(\int_T g_t d\mu g_t^*\right) \operatorname{tr}\left(\int_T F_t d\mu F_t^*\right).$$

对 $G\in L_2(\mu)^l$, 记 $G^*\triangleq(g_1^*, \cdots, g_l^*)$. 于是由上式知

$$\begin{aligned}\left\|\int_T F_t d\mu G_t^*\right\|^2 &= \sum_{j=1}^{l}\left\|\int_T F_t d\mu g_{jt}^*\right\|^2 \\ &\leqslant \sum_{j=1}^{l}\left(\int_T g_{jt} d\mu g_{jt}^*\right)\operatorname{tr}\left(\int_T F_t d\mu F_t^*\right) \\ &= \operatorname{tr}\left(\int_T G_t d\mu G_t^*\right)\operatorname{tr}\left(\int_T F_t d\mu F_t^*\right).\end{aligned}$$

为证(5.1.2)式,只须表

$$\begin{aligned}\operatorname{tr}\left(\int_T F_t d\mu F_t^*\right) &= \int_T \operatorname{tr}\left(F_t\left(\frac{d\mu}{d\operatorname{tr}\mu}\right)_t F_t^*\right) d\operatorname{tr}\mu \\ &= \int_T \left\|F_t\left(\frac{d\mu}{d\operatorname{tr}\mu}\right)_t^{\frac{1}{2}}\right\|^2 d\operatorname{tr}\mu,\end{aligned}$$

再根据普通积分的 Cauchy 不等式即可得到. ▌

§ 5.2 二阶可测过程对阵测度的积分

定义 5.2.1 设 z_T 为 m 阶 σ 有限阵测度空间 $(T, \mathscr{T}, \mu)$ 上的 m 维二阶可测过程. 若下式成立:

$$\operatorname{tr}\left(\int_T R(z_t) d\mu\right) < \infty, \tag{5.2.1}$$

则称 z_T 为(对 μ)二阶可积过程. 若存在 $(T_n)\subset\mathscr{T}$, $T_n\uparrow T$, 使

$$\operatorname{tr}\left(\int_{T_n} R(z_t) d\mu\right) < \infty \quad (n\geqslant 1),$$

则称 z_T 为(对 μ)二阶 σ 可积过程.

定理 5.2.2 设 z_T 为 $(T, \mathscr{T}, \mu)$ 上的 m 维二阶可积过程, 则 $\mathscr{R}(z_T)^k\subset L_2(\mu)^k$, 且当 $(F_n)\subset\mathscr{R}(z_T)^k$, $F\in\mathscr{R}(z_T)^k$ 及 $S_z(F_n-F)\to 0$ 时, 有 $F_n \xrightarrow{L_2(\mu)} F$. 又对任一 $H\in L_2(\mu)^k$, 令

$$\widehat{H}_t \triangleq \int_T H_s d\mu_s R(z_s, z_t) = \int_T H_s d\mu_s (Z_s^{z_t})^*, \quad (t\in T), \tag{5.2.2}$$

则 $\widehat{H}\in\mathscr{R}(z_T)^k$，且对任一 $G\in\mathscr{R}(z_T)^l$ 有

$$S_z(\widehat{H},\ G)=\int_T H_t\,d\mu G_t^*. \tag{5.2.3}$$

证 设 $F\in\mathscr{R}(z_T)^k$. 由定理 5.1.2, F 为可测. 再由引理 1.1.2,

$$\begin{aligned}
&\operatorname{tr}\left(\int_T F_t\,d\mu F_t^*\right)\\
&=\operatorname{tr}\left[\int_T R(z(F),\ z_t)\left(\frac{d\mu}{d\operatorname{tr}\mu}\right)_t R(z(F),\ z_t)^*\,d\operatorname{tr}\mu\right]\\
&=\int_T\left\|R\left(z(F),\left(\frac{d\mu}{d\operatorname{tr}\mu}\right)_t^{\frac{1}{2}}z_t\right)\right\|^2 d\operatorname{tr}\mu\\
&\leqslant\operatorname{tr}S_z(F)\int_T\operatorname{tr}\left[R(z_t)\left(\frac{d\mu}{d\operatorname{tr}\mu}\right)_t\right]d\operatorname{tr}\mu\\
&=\operatorname{tr}S_z(F)\operatorname{tr}\left(\int_T R(z_t)\,d\mu\right)<\infty.
\end{aligned}$$

故 $F\in L_2(\mu)^k$. 在上式中用 F_n-F 代替 F，即可推出 $S_z(F_n-F)\to 0$ 蕴含 $F_n\xrightarrow{L_2(\mu)}F$.

又设 $H\in L_2(\mu)^k$，且由(5.2.2)定义 $\widehat{H}$（由上述已证的事实，知 $Z^{z_t}\in\mathscr{R}(z_T)^m\subset L_2(\mu)^m$，所以 $\widehat{H}$ 是有意义的）. 于是，由(5.1.1′)和引理 1.1.2，对任意有限集 $\{t_j\}\subset T$, $\{c_j\}\subset C^m$,

$$\begin{aligned}
&\left\|\sum_j\widehat{H}_{t_j}c_j\right\|^2=\left\|\int_T H_t\,d\mu R\left(z_t,\ \sum_j c_j^*z_{t_j}\right)\right\|^2\\
&\leqslant\operatorname{tr}\left(\int_T H_t\,d\mu H_t^*\right)\\
&\quad\times\operatorname{tr}\left[\int_T R\left(\sum_j c_j^*z_{t_j},\ z_t\right)d\mu R\left(z_t,\ \sum_j c_j^*z_{t_j}\right)\right]\\
&=\operatorname{tr}\left(\int_T H_t\,d\mu H_t^*\right)\left(\int_T\left\|R\left(\sum_j c_j^*z_{t_j},\ \left(\frac{d\mu}{d\operatorname{tr}\mu}\right)^{\frac{1}{2}}z_t\right)\right\|^2 d\operatorname{tr}\mu\right.\\
&\leqslant\operatorname{tr}\left(\int_T H_t\,d\mu H_t^*\right)\operatorname{tr}\left(\int_T R(z_t)\,d\mu\right)R\left(\sum_j c_j^*z_{t_j}\right).
\end{aligned}$$

故由定理 1.2.9 知 $\widehat{H}\in\mathscr{R}(z_T)^k$.

最后来证明(5.2.3)式. 首先，设 $x=\sum_j c_j^*z_{t_j}\in\mathscr{H}_0(z_T)$，则

$$S_z(\widehat{H}, Z^x)=R(z(\widehat{H}), \sum_j c_j^* z_{t_j})=\sum_j \widehat{H}_{t_j} c_j$$

$$=\int_T H_t d\mu\, R\Big(z_t, \sum_j c_j^* z_{t_j}\Big)=\int_T H_t d\mu (Z_t^x)^*.$$

其次，设 $x\in\mathscr{H}(z_T)$，则存在$(x_n)\subset\mathscr{H}_0(z_T)$使 $R(x_n-x)\to 0$. 故有 $S_z(Z^{x_n}-Z^x)\to 0$，因此由已证部份 $Z^{x_n}\xrightarrow{L_2(\mu)} Z^x$. 在等式

$$S_z(\widehat{H}, Z^{x_n})=\int_T H_t d\mu (Z_t^{x_n})^*$$

两边取极限. 注意由定理1.2.7, vi)有

$$\|S_z(\widehat{H}, Z^{x_n})-S_z(\widehat{H}, Z^x)\|^2$$
$$\leqslant \operatorname{tr} S_z(\widehat{H}) \operatorname{tr} S_z(Z^{x_n}-Z^x)\to 0,$$

由(5.1.2)式有

$$\left\|\int_T H_t d\mu (Z_t^{x_n})^*-\int_T H_t d\mu (Z_t^x)^*\right\|$$
$$\leqslant \operatorname{tr}\Big(\int_T H_t d\mu H_t^*\Big) \operatorname{tr}\Big[\int_T (Z_t^{x_n}-Z_t^x) d\mu (Z_t^{x_n}-Z_t^x)^*\Big]$$
$$\to 0,$$

即知(5.2.3)式对任意的 $G=Z^x\in\mathscr{R}(z_T)$成立，从而对任意的 $G\in\mathscr{R}(z_T)^k$也成立. ∎

定义 5.2.3　设 z_T 为m阶 σ 有限阵测度空间 $(T, \mathscr{T}, \mu)$ 上的m维二阶可积过程. 对任一 $H\in L_2(\mu)^k$, 按(5.2.2)式给出 $\widehat{H}$, 则定义随机积分

$$\int_T H_t d\mu\, z_t \triangleq z(\widehat{H}). \tag{5.2.4}$$

其中 $z(\widehat{H})$为 $\widehat{H}$ 在 $\mathscr{H}(z_T)^k$ 中的逆再生表示.

由定理 5.2.2, $\widehat{H}\in\mathscr{R}(z_T)^k$, 所以这一定义是唯一确定的.

下一定理给出了随机积分(5.2.4)的基本性质，它表明了该定义的"合理性". 实际上，以后我们要利用的就只是积分的这些性质.

定理 5.2.4　由(5.2.4) 式定义的随机积分有如下基本性质:

i) $E\Big(\int_T H_t d\mu\, z_t\Big)=\int_T H_t d\mu (Ez_t).$

ii) 对任意的 $y\in\mathscr{L}_2^1$ 有

$$R\left(\int_T H_t d\mu\, z_t,\ y\right)=\int_T H_t d\mu\, R(z_t,\ y).$$

iii) 对任意的 $H\in L_2(\mu)^k$, $J\in L_2(\mu)^l$, $C\in C^{p\times k}$, $D\in C^{p\times l}$ 有

$$\int_T (CH_t+DJ_t)d\mu\, z_t=C\int_T H_t d\mu\, z_t+D\int_T J_t d\mu\, z_t.$$

iv) 有不等式

$$\left\|R\left(\int_T H_t d\mu\, z_t,\ \int_T J_t d\mu\, z_t\right)\right\|^2$$

$$\leqslant \operatorname{tr}\left(\int_T H_t d\mu\, H_t^*\right)\operatorname{tr}\left(\int_T J_t d\mu\, J_t^*\right)\left[\operatorname{tr}\left(\int_T R(z_t)d\mu\right)\right]^2$$

成立. 特别有

$$\left\|R\left(\int_T H_t d\mu\, z_t\right)\right\|\leqslant \operatorname{tr}\left(\int_T H_t d\mu\, H_t^*\right)\operatorname{tr}\left(\int_T R(z_t)d\mu\right).$$

v) 若 $H_n\xrightarrow{L_2(\mu)}H$, 则

$$\int_T H_{nt} d\mu\, z_t\xrightarrow{\mathscr{L}_2}\int_T H_t d\mu\, z_t.$$

vi) 若 z_T, z_{nT} 都是$(T,\ \mathscr{T},\ \mu)$上的m维二阶可积过程, 且 $Ez_{nT}^*\xrightarrow{L_2(\mu)}Ez_T^*$, $\operatorname{tr}\left(\int_T R(z_{nt}-z_t)d\mu\right)\to 0$, 则

$$\int_T H_t d\mu\, z_{nt}\xrightarrow{\mathscr{L}_2}\int_T H_t d\mu\, z_t.$$

vii) 设 $S\in\mathscr{T}$, 1_S 为其示性函数. 若 $\operatorname{tr}\left(\int_S H_t d\mu\, H_t^*\right)<\infty$, 则

$$\int_T (1_S H)_t d\mu\, z_t=\int_S H_t d\mu\, z_t\in\mathscr{H}(z_S)^k.$$

viii) $\displaystyle\int_T H_t d\mu\, z_t=\int_T H_t\left(\frac{d\mu}{d\operatorname{tr}\mu}\right)_t z_t d\operatorname{tr}\mu.$

证　i) 记 $Ez_t\triangleq f_t^*$, $t\in T$. 由 z_T 的正则性假设即知 $f\in\mathscr{R}(z_T)$. 再由定理5.2.2知 $f\in L_2(\mu)$. 于是由定理1.2.13和(5.2.3)式即得

$$E\left(\int_T H_t d\mu z_t\right)=Ez(\widehat{H})=S_s(\widehat{H},\ f)=\int_T H_t d\mu f_t^*$$

$$=\int_T H_t d\mu(Ez_t).$$

ii) 设 $y\in\mathscr{L}_2^1$. 则由定理 2.1.2 和(5.2.3),

$$R\left(\int_T H_t d\mu z_t,\ y\right)=R(z(\widehat{H}),\ y)=R(z(\widehat{H}),\ \pi(y|z_T))$$

$$=S_s(\widehat{H},\ Z^y)=\int_T H_t d\mu(Z_t^y)^*$$

$$=\int_T H_t d\mu R(z_t,\ y).$$

iii) 由 $z(\widehat{CH+DJ})=z(C\widehat{H}+D\widehat{J})=Cz(\widehat{H})+Dz(\widehat{J})$ 即得.

iv) 由 ii)与(5.1.1′)以及引理 1.1.2:

$$\left\|R\left(\int_T H_t d\mu z_t, \int_T J_t d\mu z_t\right)\right\|^2$$

$$=\left\|\int_T H_t d\mu_t R\left(z_t, \int_T J_s d\mu_s z_s\right)\right\|^2$$

$$\leqslant \mathrm{tr}\left(\int_T H_t d\mu H_t^*\right)$$

$$\times\int_T\left\|\int_T J_s d\mu_s R\left(z_s, \left(\frac{d\mu}{d\,\mathrm{tr}\mu}\right)_t^{\frac{1}{2}} z_t\right)\right\|^2 d_t\mathrm{tr}\,\mu_t$$

$$\leqslant \mathrm{tr}\left(\int_T H_t d\mu H_t^*\right)\mathrm{tr}\left(\int_T J_s d\mu J_s^*\right)$$

$$\times\int_T\int_T\left\|R\left(\left(\frac{d\mu}{d\,\mathrm{tr}\mu}\right)_t^{\frac{1}{2}} z_t, \left(\frac{d\mu}{d\,\mathrm{tr}\mu}\right)_s^{\frac{1}{2}} z_s\right)\right\|^2 d\,\mathrm{tr}\,\mu_s d\,\mathrm{tr}\,\mu_t$$

$$\leqslant \mathrm{tr}\left(\int_T H_t d\mu H_t^*\right)\mathrm{tr}\left(\int_T J_t d\mu J_t^*\right)\left[\mathrm{tr}\left(\int_T R(z_t)d\mu\right)\right]^2.$$

v) 由 iv):

$$\left\|R\left(\int_T H_{nt} d\mu z_t-\int_T H_t d\mu z_t\right)\right\|$$

$$\leqslant \mathrm{tr}\left[\int_T (H_{nt}-H_t)d\mu(H_{nt}-H_t)^*\right]\mathrm{tr}\left(\int_T R(z_t)d\mu\right).$$

又注意在 $\mathscr{H}(z_T)$ 中，$\mathscr{L}_2$ 收敛(即以范数 R^+ 收敛)等价于以范数 R 收敛．由此即得所需结论．

vi) 由 iv)：

$$\left\| R\left(\int_T H_t d\mu\, z_{nt} - \int_T H_t d\mu\, z_t\right)\right\|$$
$$\leqslant \operatorname{tr}\left(\int_T H_t d\mu\, H_t^*\right) \operatorname{tr}\left(\int_T R(z_{nt} - z_t) d\mu\right),$$

又由(5.1.1′)：

$$\left\| \int_T H_t d\mu (Ez_{nt}) - \int_T H_t d\mu (Ez_t)\right\|^2$$
$$\leqslant \operatorname{tr}\left(\int_T H_t d\mu\, H_t^*\right)\left[\int_T (Ez_{nt} - Ez_t)^* d\mu (Ez_{nt} - Ez_t)\right].$$

由此即得所需结论．

vii) 由于 $\int_S H_t d\mu\, z_t \in \mathscr{H}(z_S)^k \subset \mathscr{H}(z_T)^k$，且对任意 $s \in T$ 有

$$R\left(\int_S H_t d\mu\, z_t,\ z_s\right) = \int_S H_t d\mu_t R(z_t,\ z_s)$$
$$= \int_T (1_S H)_t d\mu_t R(z_t, z_s) = \widehat{(1_S H)}_s,$$

故得 $$\int_S H_t d\mu\, z_t = z\left(\widehat{1_S H}\right) = \int_T (1_S H)_t d\mu\, z_t.$$

viii) 由于

$$\widehat{H}_s = \int_T H_t \left(\frac{d\mu}{d\operatorname{tr}\mu}\right)_t R(z_t,\ z_s)\, d\operatorname{tr}\mu_t$$
$$= R\left(\int_T H_t \left(\frac{d\mu}{d\operatorname{tr}\mu}\right)_t z_t\, d\operatorname{tr}\mu,\ z_s\right). \quad ∎$$

设 F 为 T 上的 $C^{k\times m}$ 值函数．若 F 使积分方程

$$F_t = \int_T \tilde{F}_s d\mu_s R(z_s,\ z_t),\ t \in T$$

有解 $\tilde{F} \in L_2(\mu)^k$，则由(5.2.2)、(5.2.3)及(5.2.4)式即可知道，$F \in \mathscr{R}(z_T)^k$，对任一 $G \in \mathscr{R}(z_T)^l$ 有 $S_z(F,\ G) = \int_T \tilde{F}_t d\mu\, \tilde{G}_t^*$，及

$z(F)=\int_T \widetilde{F}_t d\mu z_t$. 其中后者虽然是作为随机积分的定义，但由于有定理 5.2.4, 它可用求和法进行近似计算. 故对于牵涉到这样的 F 的线性统计问题，只要上述积分方程能求解，便可以得到基本满意的解决(参见第二章最后的说明). 这类积分方程通常称为 Wiener-Hopf 方程. 当 $T=(-\infty, t_0]$ 或其中的整数点集，且 z_T 为二阶平稳过程(即宽平稳过程)时，Wiener 曾经深刻地研究过这类积分方程的解，并用以解决平稳过程的线性预测与滤波问题. 但值得注意的是，这样的 F 一般并未穷尽整个 $\mathscr{R}(z_T)^k$; 相应地，全体随机积分 $\left\{\int_T H_t d\mu z_t;\ H\in \boldsymbol{L}_2(\mu)^k\right\}$ 也未穷尽 $\mathscr{H}(z_T)^k$, 即未穷尽由 z_T 生成的全部 k 维线性统计量.

§5.3 平方可积函数对正交随机测度的积分

定义 5.3.1 设 $(\boldsymbol{T}, \mathscr{T}, \mu)$ 为 m 阶 σ 有限阵测度空间，$\mathscr{S}$ 为生成 $\mathscr{T}$ 的一个 π 类(即 $\mathscr{S}$ 对有限交封闭，且 $\sigma(\mathscr{S})=\mathscr{T}$), $\mathscr{S}$ 的元都有有限 μ 测度，且存在 $(\boldsymbol{T}_n)\subset\mathscr{S}$, 使 $\boldsymbol{T}_n\uparrow\boldsymbol{T}$. 又设 $\boldsymbol{w}_{\mathscr{S}}\triangleq(\boldsymbol{w}(S), S\in\mathscr{S})$ 为以 $\mathscr{S}$ 为指标集的 m 阶正则随机过程(随机集函数). 若对任意 S_1、$S_2\in\mathscr{S}$ 都有

$$R(\boldsymbol{w}(S_1),\ \boldsymbol{w}(S_2))=\mu(S_1\cap S_2)$$

成立. 则称 $\boldsymbol{w}_{\mathscr{S}}$ 为 $\mathscr{S}$ 上以 μ 为核的 m 维正交随机测度. 这一定义是通常实轴上的正交增量过程的推广. 详细叙述参见第六章.

以下恒设 $\mathscr{S}$ 满足定义规定的条件.

对任一 $H\in\boldsymbol{L}_2(\mu)^k$, 定义 $\mathscr{S}$ 上的 $k\times m$ 阵函数 H^μ 为:

$$H^\mu(S)\triangleq\int_S H_t d\mu=\int_T H_t d\mu I^*_{St},\ S\in\mathscr{S}, \qquad (5.3.1)$$

其中 $I_S\triangleq 1_S I_m$, I_m 为 m 阶单位阵，1_S 为 S 的示性函数. 又记

$$\frac{dH^\mu}{d\mu}\triangleq H \qquad (5.3.2)$$

下面的引理表明了 $\frac{dH^\mu}{d\mu}$ 在 $L_2(\mu)^k$ 中被 H^μ 唯一确定.

引理 5.3.2 设 $H, J\in L_2(\mu)^k$, 则 $H^\mu=J^\mu$ 的必要充分条件是 $H=J$, a.e. μ.

证 充分性是显然的. 必要性等价于: 若对任意 $S\in\mathscr{S}$ 有 $\int_S H_t d\mu=0$, 则对任意 $S\in\mathscr{T}$ 也有 $\int_S H_t d\mu=0$. 根据对 $\mathscr{S}$ 所作的假设及积分的性质, 子集类

$$\mathscr{C}\triangleq\left\{S\in\mathscr{T}:\int_S H_t d\mu=0\right\}$$

是一包含 π 类 $\mathscr{S}$ 的 λ 类(即满足 $T\in\mathscr{C}$; S_1、$S_2\in\mathscr{S}$ 及 $S_1\subset S_2\Rightarrow S_2\backslash S_1\in\mathscr{C}$; $(S_n)\subset\mathscr{C}$ 及 $S_n\uparrow\Rightarrow\bigcup_n S_n\in\mathscr{C}$). 因而由单调类定理得 $\mathscr{C}=\sigma(\mathscr{S})=\mathscr{T}$. ∎

定理 5.3.3 设 $w_\mathscr{S}$ 是以 μ 为核的 m 维正交随机测度. 则其再生核表示空间 $(\mathscr{R}(w_\mathscr{S}), S_w)$ 为

$$\mathscr{R}(w_\mathscr{S})^k=\{H^\mu: H\in L_2(\mu)^k\},$$

$$S_w(H^\mu, J^\mu)=\int_T H_t d\mu J_t^*,\ H\in L_2(\mu)^k,\ J\in L_2(\mu)^l.$$

证 只须对 $k=l=1$ 的情形加以证明. 令

$$\mathscr{R}\triangleq\{h^\mu: h\in L_2(\mu)\},$$

$$S(h_1^\mu, h_2^\mu)\triangleq\int_T h_{1t}d\mu h_{2t}^*,\ (h_1、h_2\in L_2(\mu)),$$

则由引理 5.3.2 知 $(\mathscr{R}, S)$ 为一与 $L_2(\mu)$ 同构的 Hilbert 空间.

由(5.1.1′), 对任意有限集 $\{S_j\}\subset\mathscr{S}$, $\{c_j\}\subset C^m$ 有

$$\begin{aligned}\left|\sum_j h^\mu(S_j)c_j\right|^2&=\left|\int_T h_t d\mu\left(\sum_j c_j^* I_{S_j t}\right)^*\right|^2\\&\leqslant\left(\int_T h_t d\mu h_t^*\right)\\&\quad\times\left(\int_T\left(\sum_j c_j^* I_{S_j t}\right)d\mu\left(\sum_j c_j^* I_{S_j t}\right)^*\right)\\&=\left(\int_T h_t d\mu h_t^*\right)\left(\sum_{i,j}c_i^*\mu(S_i\cap S_j)c_j\right)\end{aligned}$$

$$=\left(\int_T h_t d\mu\, h_t^*\right)R\left(\sum_j c_j^* w(S_j)\right),$$

故由定理1.2.9, $h^\mu\in\mathscr{R}(w_\mathscr{S})$. 于是 $\mathscr{R}\subset\mathscr{R}(w_\mathscr{S})$.

又设 $x=\sum_j c_j^* w(S_j)\in\mathscr{H}_0(w_\mathscr{S})$. 按(1.2.2)式规定的记法, x 在 $\mathscr{R}_0(w_\mathscr{S})$中的再生表示记作 W^x, 而对任意的 $S\in\mathscr{S}$,

$$\begin{aligned}W^x(S)&\triangleq R(x,\ w(S))=\sum_j c_j^* R(w(S_j),\ w(S))\\&=\sum_j c_j^*\mu(S_j\cap S)=\int_S\left(\sum_j c_j^* I_{S_j}\right)_t d\mu\\&=\left(\sum_j c_j^* I_{S_j}\right)^\mu(S).\end{aligned}$$

故 $W^x\in\mathscr{R}$, 且 $\dfrac{dW^x}{d\mu}=\sum_j c_j^* I_{S_j}$, 从而 $\mathscr{R}_0(w_\mathscr{S})\subset\mathscr{R}$. 且由再生性知

$$\begin{aligned}S_w(h^\mu,\ W^x)&=S_w\left(h^\mu,\ \sum_j c_j^* W^{w(S_j)}\right)=\sum_j S_w(h^\mu,\ W^{w(S_j)})c_j\\&=\sum_j h^\mu(S_j)\,c_j=\int_T h_t\,d\mu\left(\sum_j c_j^* I_{S_j}\right)_t^*\\&=S(h^\mu,\ W^x).\end{aligned}$$

由此即得$(\mathscr{R}(w_\mathscr{S}),\ S_w)=(\mathscr{R},\ S)$. ∎

定义 5.3.4 设$(T,\mathscr{T},\mu)$为m阶σ有限阵测度空间, $\mathscr{S}(\subset\mathscr{T})$满足定义5.3.1规定的条件, $w_\mathscr{S}$为以μ为核的m维正交随机测度, $H\in L_2(\mu)^k$. 则定义随机积分

$$\int_T H_t\,dw\triangleq w(H^\mu).\tag{5.3.3}$$

其中 $w(H^\mu)$为 H^μ 在 $\mathscr{H}(w_\mathscr{S})^k$中的逆再生表示(按定义1.2.6规定的记号).

由定理5.3.3, $H^\mu\in\mathscr{R}(w_\mathscr{S})^k$, 所以这一定义是唯一确定的.

定理 5.3.5 由(5.3.3)定义的随机积分有如下基本性质: 设 H, $H_n\in L_2(\mu)^k$, $J\in L_2(\mu)^l$, $C\in C^{p\times k}$, $D\in C^{p\times l}$, 则

i) 存在 $f\in L_2(\mu)$, 使 $Ew(S)^*=\int_S f\,d\mu$, $S\in\mathscr{S}$, 且

$$E\left(\int_T H_t\,dw\right)=\int_T H_t\,d\mu\,f^*.$$

ii) $R\left(\int_T H_t dw, \int_T J_t dw\right)=\int_T H_t d\mu J_t^*$.

iii) $\int_T H_t dw=\int_T J_t dw$ 的必要充分条件是 $H=J$, a.e. μ.

iv) $\int_T (CH_t+DJ_t)dw=C\int_T H_t dw+D\int_T J_t dw$.

v) $\int_T H_{nt} dw \xrightarrow{\mathscr{L}_2} \int_T H_t dw$ 的必要充分条件是 $H_n \xrightarrow{L_2(\mu)} H$.

vi) 对任一 $S\in\mathscr{S}$, $\int_T I_{St} dw=w(S)$.

vii) $\mathscr{H}(w_{\mathscr{S}})^k=\left\{\int_T H_t dw;\ H\in L_2(\mu)^k\right\}$.

证 以上诸性质都可由逆再生表示的性质及定理 1.2.11 和 1.2.12 以及引理 5.3.2 和定理 5.3.3 直接推出. ∎

系 5.3.6 由 5.3.4 定义的随机积分在下述意义下与 $\mathscr{S}$ 的选取无关: 设 $\mathscr{S}_1$ 为满足定义 5.3.1 规定的任一 π 类, 令

$$w_{\mathscr{S}_1}\triangleq\left\{w(S_1)\triangleq\int_T I_{S_1t} dw,\ S_1\in\mathscr{S}_1\right\}, \qquad (5.3.4)$$

则 $w_{\mathscr{S}_1}$ 是 $\mathscr{S}_1$ 上以 μ 为核的 m 维正交随机测度, 且由 $w_{\mathscr{S}_1}$ 按定义 5.3.4 给出的随机积分与由 $w_{\mathscr{S}}$ 给出的完全相同; 从而又有 $\mathscr{H}(w_{\mathscr{S}})=\mathscr{H}(w_{\mathscr{S}_1})$.

证 由 ii), $w_{\mathscr{S}_1}$ 是以 μ 为核的正交随机测度. 设 $H\in L_2(\mu)^k$. 我们暂用 $\int_T^{(1)} H_t dw$ 表示由 $w_{\mathscr{S}_1}$ 定义的随机积分. 因为 $\mathscr{H}(w_{\mathscr{S}_1})\subset\mathscr{H}(w_{\mathscr{S}})$, 由 vii), 必存在 $H^{(1)}\in L_2(\mu)^k$ 使 $\int_T^{(1)} H_t dw=\int_T H_t^{(1)} dw$. 但由定义及 ii), 对任意的 $S_1\in\mathscr{S}_1$ 有

$$\int_{S_1} H_t d\mu=H^\mu(S_1)=R\left(\int_T^{(1)} H_t dw,\ w(S_1)\right)$$

$$=R\left(\int_T H_t^{(1)} dw,\ \int_T I_{S_1t} dw\right)=\int_{S_1} H_t^{(1)} dw,$$

于是由引理 5.3.2 即知 $H=H^{(1)}$, a.e. μ. ∎

由这个系还知道, 我们总可以通过(5.3.4)将一个 $\mathscr{S}$ 上的正

交随机测度 $\boldsymbol{w}_{\mathscr{T}}$ 扩充到有有限 μ 测度的 $S\in\mathscr{T}$ 全体上，它们显然组成满足定义 5.3.1 规定的最大 π 类. 为使符号简便起见，我们把这个扩充的正交随机测度就记作 $\boldsymbol{w}_{\mathscr{T}}$. 由 $\boldsymbol{w}_{\mathscr{T}}$ 与 $\boldsymbol{w}_{\mathscr{S}}$ 定义的随机积分是相同的，而且 $\mathscr{H}(\boldsymbol{w}_{\mathscr{T}})=\mathscr{H}(\boldsymbol{w}_{\mathscr{S}})$.

下面的定理及其系分别表明了随机积分(5.3.3)的代换原则与局部性质.

定理 5.3.7 设$(\boldsymbol{T},\ \mathscr{T},\ \mu)$为 m 阶 σ 有限阵测度空间，$\boldsymbol{w}_{\mathscr{T}}$ 是以 μ 为核的 m 维正交随机测度，$G\in\sigma\boldsymbol{L}_2(\mu)^l$. 令

$$\nu(S)\triangleq\int_S G_t\,d\mu\,G_t^*,\ S\in\mathscr{T},$$

$$\mathscr{S}\triangleq\left\{S\in\mathscr{T}:\ \operatorname{tr}\left(\int_S G_t\,d\mu\,G_t^*\right)<\infty\right\},$$

$$(G\cdot\boldsymbol{w})_{\mathscr{S}}\triangleq\left\{(G\cdot\boldsymbol{w})(S)\triangleq\int_T(1_S G)_t\,d\boldsymbol{w},\ S\in\mathscr{S}\right\}.$$

则 ν 是$(\boldsymbol{T},\ \mathscr{T})$上的 l 阶 σ 有限阵测度，$\mathscr{S}$ 是满足定义 5.3.1 规定的 π 类，$(G\cdot\boldsymbol{w})_{\mathscr{S}}$ 是以 ν 为核的 l 维正交随机测度.

设 H 为$(\boldsymbol{T},\ \mathscr{T})$上的可测 $\boldsymbol{C}^{k\times l}$ 值函数. 则 $H\in\boldsymbol{L}_2(\nu)^k$ 的必要充分条件是 $HG\in\boldsymbol{L}_2(\mu)^k$. 此时有下式成立：

$$\int_T H_t\,d(G\cdot\boldsymbol{w})=\int_T(HG)_t\,d\boldsymbol{w}.$$

证 定理的前一部分由普通积分的性质及定理 5.3.5 的 ii) 立即推出.

至于后一部分，因为可表为

$$\nu(S)=\int_S G_t\left(\frac{d\mu}{d\operatorname{tr}\mu}\right)_t G_t^*\,d\operatorname{tr}\mu,\ S\in\mathscr{T}.$$

由 Radon-Nikodym 导数的性质即知

$$\int_T H_t\,d\nu\,H_t^*=\int_T H_t G_t\left(\frac{d\mu}{d\operatorname{tr}\mu}\right)_t G_t^* H_t^*\,d\operatorname{tr}\mu$$

$$=\int_T(HG)_t\,d\mu(HG)_t^*.$$

而此即说明了 $H\in\boldsymbol{L}_2(\nu)^k$ 当且仅当 $HG\in\boldsymbol{L}_2(\mu)^k$. 对这样的 H，存在 $\boldsymbol{L}_2(\nu)^k$ 中的简单函数列(H_n)，使 $H_n\xrightarrow{\boldsymbol{L}_2(\nu)}H$，因而

$H_nG \xrightarrow{L_2(\mu)} HG$. 又由 5.3.5 之 vi), 对任意 $S\in\mathscr{S}$ 有

$$\int_T I_{St}d(G.\boldsymbol{w})=(G\cdot\boldsymbol{w})\ (\boldsymbol{S})=\int_T (I_SG)_t d\boldsymbol{w}.$$

于是由 5.3.5 的 iv),

$$\int_T H_{nt}d(G.\boldsymbol{w})=\int_T (H_nG)_t d\boldsymbol{w}.$$

令 $n\to\infty$, 根据 5.3.5 的 v), 即得

$$\int_T H_t d(G\cdot\boldsymbol{w})=\int_T (HG)_t d\boldsymbol{w}. \quad \blacksquare$$

系 5.3.8 设 $(\boldsymbol{T},\ \mathscr{T},\ \mu)$ 为 m 阶 σ 有限阵测度空间, $\boldsymbol{w}_{\mathscr{T}}$ 是以 μ 为核的正交随机测度, $\boldsymbol{T}_0\in\mathscr{T}$. 令

$$\mu^0(\boldsymbol{S})\triangleq\mu(\boldsymbol{S}\cap\boldsymbol{T}_0),\ \boldsymbol{S}\in\mathscr{T},$$

$$\boldsymbol{w}^0_{\mathscr{T}}\triangleq\{\boldsymbol{w}^0(\boldsymbol{S})\triangleq\boldsymbol{w}(\boldsymbol{S}\cap\boldsymbol{T}_0);\ \boldsymbol{S}\in\mathscr{T}\ 且\ \mathrm{tr}\,\mu(\boldsymbol{S})<\infty\}.$$

则 μ^0 是 $(\boldsymbol{T},\ \mathscr{T})$ 上的 m 阶 σ 有限阵测度, $\boldsymbol{w}^0_{\mathscr{T}}$ 是以 μ^0 为核的 m 维正交随机测度. $H\in\boldsymbol{L}_2(\mu^0)^k$ 的必要充分条件是 $1_{T_0}H\in\boldsymbol{L}_2(\mu)^k$, 即 $\mathrm{tr}\left(\int_{T_0} H_t d\mu H_t^*\right)<\infty$. 此时有下式成立:

$$\int_{T_0} H_t d\boldsymbol{w}\triangleq\int_T (1_{T_0}H)_t d\boldsymbol{w}=\int_T H_t d\boldsymbol{w}^0.$$

它是 $\mathscr{H}(\boldsymbol{w}^0_{\mathscr{T}})^k=\mathscr{H}(\boldsymbol{w}_{\mathscr{T}\cap T_0})^k$ 中的元.

证 只须在定理 5.3.7 中取 $G=I_{T_0}$ 即可. $\blacksquare$

§5.4 应用于对局部平方可积鞅的积分

本节的内容并不是今后所必需的. 但作为一个应用的例子却能深刻地说明我们在一般的可测空间上来定义正交随机测度及对它的积分并非"无的放矢"的单纯推广; 它可以把一些相当现代且有用的概念包括进来.

本节中关于鞅论的基本理论可参看参考文献[9].

设 $(\Omega,\mathscr{F},(\mathscr{F}_t)_{t\in[0,\infty)},P)$ 是一带有满足通常条件的上升子 σ 域族 $(\mathscr{F}_t)_{t\in[0,\infty)}$ 的概率空间, $([0,\ \infty)\times\Omega,\ \mathscr{P})$ 是与之联系的可料

可测空间，其中 $\mathscr{P}$ 是 $[0,\infty)\times\Omega$ 上的可料 σ 域，它由 π 类 $\{[\![0,\tau]\!]:\tau$ 为 $(\mathscr{F}_t)$ 停时$\}$ 所生成，其中 $[\![0,\tau]\!]$ 表示"随机区间"：

$$[\![0,\tau]\!]\triangleq\{(t,\omega):0\leqslant t\leqslant\tau(\omega),\text{当 }\tau(\omega)<\infty;\quad 0\leqslant t<\infty,\text{当 }\tau(\omega)=\infty\}.$$

设 $x_{[0,\infty)}$ 为一 m 维右连续零初值局部平方可积 $(\mathscr{F}_t)$ 鞅，即 $x_0=0$，且存在 $(\mathscr{F}_t)$ 停时列 (τ_n)，$\tau_n\uparrow\infty$，使每一 $x^{\tau_n}_{[0,\infty)}\triangleq(x_{t\wedge\tau_n},t\in[0,\infty))$ 为 m 维右连续平方可积 $(\mathscr{F}_t)$ 鞅. 于是根据著名的 Doob–Meyer 分解定理，存在唯一的 m 阶 Hermite 阵(实对称阵)值、右连续、零初值局部可积的可料增过程 $\langle x\rangle_{[0,\infty)}$，使 $(xx^*)_{[0,\infty)}-\langle x\rangle_{[0,\infty)}$ 为 m 阶矩阵值局部鞅. 我们称 $\langle x\rangle_{[0,\infty)}$ 为 $x_{[0,\infty)}$ 的二次变差过程；实际上，$\langle x\rangle_{[0,\infty)}$ 的对角线元就是 $x_{[0,\infty)}$ 的各相应分量的二次变差，而其非对角线元则是相应两个分量间的互二次变差.

现在，对任意停时 τ 定义 $\mu_x([\![0,\tau]\!])\triangleq E\langle x\rangle_\tau$. 则 μ_x 可唯一地扩张成 $([0,\infty)\times\Omega,\mathscr{P})$ 上的 m 阶 σ 有限阵测度 $(\mu_x([\![0,\tau_n]\!])$ 为有限值)，通常称为由 $\langle x\rangle_{[0,\infty)}$ 导出的 Dolean 测度. 又令

$$\mathscr{S}\triangleq\{[\![0,\tau\wedge\tau_n]\!]:n\in N,\ \tau\text{ 为任意停时}\},$$

$$x([\![0,\tau\wedge\tau_n]\!])\triangleq x_{\tau\wedge\tau_n},\ \tau\text{ 为任意停时},$$

则 $\mathscr{S}$ 为一 π 类，$\sigma(\mathscr{S})=\mathscr{P}$，$\mu_x$ 在 $\mathscr{S}$ 上取有限值. 又由 Doob 随机可选定理，对任意 l、$n\in N$，任意停时 ρ、τ 有

$$Ex([\![0,\tau\wedge\tau_n]\!])=Ex_{\tau\wedge\tau_n}=Ex_0=0,$$

$$\begin{aligned}R(x([\![0,\rho\wedge\tau_l]\!]),x([\![0,\tau\wedge\tau_n]\!]))&=E(x_{\rho\wedge\tau_l}x^*_{\tau\wedge\tau_n})\\&=E(x_{\rho\wedge\tau_l\wedge\tau\wedge\tau_n}x^*_{\tau\wedge\tau_n})=E(xx^*)_{\rho\wedge\tau_l\wedge\tau\wedge\tau_n}\\&=E\langle x\rangle_{\rho\wedge\tau_l\wedge\tau\wedge\tau_n}=\mu_x([\![0,\rho\wedge\tau_l]\!]\cap[\![0,\tau\wedge\tau_n]\!]),\end{aligned}$$

即 $x_{\mathscr{S}}$ 是 $\mathscr{S}$ 上以 μ_x 为核的 m 维正交随机测度.

于是，根据定义 5.3.4，对于 $([0,\infty)\times\Omega,\mathscr{P},\mu_x)$ 上的任一 $C^{k\times m}$ 值平方可积函数 H，亦即 $C^{k\times m}$ 值，且对 μ_x 平方可积的可料过程 $H_{[0,\infty)}$，可以定义它对 m 维局部平方可积鞅 $x_{[0,\infty)}$ 的随机积分：

$$\int_{[0,\infty)}H_t\,dx_t\triangleq x(H^{\mu_x}).\tag{5.4.1}$$

其中 $x(H^{\mu_x})$ 为 H^{μ_x} 在 $\mathscr{H}(x_{\mathscr{S}})^k$ 中的逆再生表示，H^{μ_x} 为 $\mathscr{S}$ 上

的 $C^{k\times m}$ 函数：

$$H^{\mu_x}([\![0,\ \tau\wedge\tau_n]\!])\triangleq\int_{[0,\tau\wedge\tau_n]}H_t(\omega)d\mu_x(t,\ \omega)$$

$$=E\left(\int_{[0,\tau\wedge\tau_n]}H_t d\langle x\rangle_t\right),$$

而右边括号中的积分是按过程 $\langle x\rangle_{[0,\infty)}$ 的轨道积分.

由定理 5.3.5、系 5.3.6、定理 5.3.7、系 5.3.8 等诸条，可以平行地推出对鞅的随机积分(5.4.1)的各种性质. 例如

$$E\left(\int_{[0,\infty)}H_t dx_t\right)=0,$$

$$R\left(\int_{[0,\infty)}H_t dx_t,\ \int_{[0,\infty)}J_t dx_t\right)=E\left(\int_{[0,\infty)}H_t d\langle x\rangle_t J_t^*\right).$$

§5.5 两种随机积分的次序交换及谱表示的推广

首先我们来讨论在 §5.2 和 §5.3 中分别定义的两种随机积分(定义 5.2.3 和 5.3.4)之间的联系，证明同时包含这两种积分的算式，其先后运算次序可以交换.

定理 5.5.1 设 $(\boldsymbol{T},\ \mathscr{T},\ \mu)$，$(\boldsymbol{U},\ \mathscr{U},\ \nu)$ 分别为 m 阶和 l 阶 σ 有限阵测度空间，$H\in\boldsymbol{L}_2(\mu)^k$. Φ 为乘积可测空间 $(\boldsymbol{T}\times\boldsymbol{U},\ \mathscr{T}\times\mathscr{U})$ 上的 $C^{m\times l}$ 值可测函数，且满足

$$\operatorname{tr}\left[\int_{\boldsymbol{T}}\left(\int_{\boldsymbol{U}}\Phi(t,\ u)d\nu_u\Phi^*(t,\ u)\right)d\mu_t\right]<\infty.$$

w_u 是以 ν 为核的 l 维正交随机测度. 则下式左右两端的随机积分存在且有以下等式成立：

$$\int_{\boldsymbol{T}}H_t d\mu_t\left(\int_{\boldsymbol{U}}\Phi(t,\ u)dw_u\right)=\int_{\boldsymbol{U}}\left(\int_{\boldsymbol{T}}H_t d\mu_t\Phi(t,\ u)\right)dw_u. \tag{5.5.1}$$

证 记

$$\Psi(s,\ t)\triangleq\int_{\boldsymbol{U}}\Phi(s,\ u)d\nu_u\Phi^*(t,\ u),\quad (s、t\in\boldsymbol{T}).$$

由普通积分性质，它是二元可测的 $C^{m\times m}$ 值函数. 因为

$$R\left(\int_{\boldsymbol{U}}\Phi(s,\ u)\,dw_u,\ \int_{\boldsymbol{U}}\Phi(t,\ u)dw_u\right)=\Psi(s,\ t),$$

且由定理假设,

$$\operatorname{tr}\left[\int_T R\left(\int_U \Phi(t, u)dw_u\right)d\mu_t\right]=\operatorname{tr}\left(\int_T \Psi(t, t)d\mu_t\right)<\infty.$$

故由定理 5.1.2 和定义 5.2.1, m 维过程 $\left(\int_U \Phi(t, u)dw_u,\ t\in T\right)$ 为对 μ 二阶可积的过程. 因此(5.5.1)式左端的随机积分是存在的.

另一方面, 由普通积分的 Fubini 定理及定理 5.2.4 的 iv),

$$\begin{aligned}
&\left\|\int_U\left(\int_T H_s d\mu_s \Phi(s, u)\right)d\nu_u\left(\int_T H_t d\mu_t \Phi(t, u)\right)^*\right\| \\
&=\left\|\int_T\left(\int_T H_s d\mu_s \Psi(s, t)\right)d\mu_t H_t^*\right\| \\
&=\left\|R\left(\int_T H_t d\mu_t\left(\int_U \Phi(t, u)dw_u\right)\right)\right\| \\
&\leqslant \operatorname{tr}\left(\int_T H_t d\mu_t H_t^*\right)\operatorname{tr}\left(\int_T \Psi(t, t)d\mu_t\right)<\infty.
\end{aligned}$$

因此(5.5.1)式右端的随机积分也存在.

最后, 对任一 $x\in\mathscr{L}_2$,

$$\begin{aligned}
&R\left(\int_U\left(\int_T H_t d\mu_t \Phi(t, u)\right)dw_u,\ x\right) \\
&=R\left(\int_U\left(\int_T H_t d\mu_t \Phi(t, u)\right)dw_u,\ w(W^x)\right) \\
&=\int_U\left(\int_T H_t d\mu_t \Phi(t, u)\right)d\nu_u\left(\frac{dW^x}{d\nu}\right)_u^* \\
&=\int_T H_t d\mu_t\left(\int_U \Phi(t, u)d\nu_u\left(\frac{dW^x}{d\nu}\right)_u^*\right) \\
&=\int_T H_t d\mu_t R\left(\int_U \Phi(t, u)dw_u,\ x\right) \\
&=R\left(\int_T H_t d\mu_t\left(\int_U \Phi(t, u)dw_u\right),\ x\right),
\end{aligned}$$

所以(5.5.1)式两端相等. ∎

由平稳过程的基本理论知道, 对于离散时间或连续时间且均方连续的零均值二阶平稳过程 z_T 分别有如下的谱表示:

$$z_t=\int_{[-\pi,\pi)} e^{itu}dw_u,\ t=0,\ \pm1,\ \pm2,\ \cdots,$$

$$z_t=\int_{(-\infty,\infty)} e^{itu}dw_u,\ t\in(-\infty,\ \infty).$$

其中 $w_U=(w_u,\ u\in U)$分别是 $U=[-\pi,\ \pi]$和$(-\infty,\ \infty)$ 上的均方左连续的正交增量过程. 实际上, 可以由 w_U 在 U 的 Borel 域 $\mathscr{B}(U)$上唯一确定一正交随机测度(参见第六章). 这一谱表示是全部二阶平稳过程理论的基本工具.

下面我们把这一概念加以推广:

定义 5.5.2 设 T 为任意指标集, $z_T=(z_t,\ t\in T)$ 为一 m 维二阶过程. 若存在 l 阶 σ 有限阵测度空间$(U,\ \mathscr{U},\ \nu)$上以 ν 为核的正交随机测度 $w_\mathscr{U}$, 及 U 上的 $C^{m\times l}$ 值函数族 $\Phi_T\triangleq\{\Phi_t;\ t\in T\}\subset L_2(U,\ \mathscr{U},\ \nu)^m$, 使

$$z_t=\int_U\Phi_t(u)dw_u,\ (t\in T),\tag{5.5.2}$$

则称 z_T 有一广义谱表示.

又用 $L_2(\Phi_T,\ \nu)$表 Φ_T 中阵函数的所有行向量函数张成的线性空间在 $L_2(U,\ \mathscr{U},\ \nu)\triangleq L_2(\nu)$中的闭包(闭子空间), 即

$$L_2(\Phi_T,\ \nu)\triangleq\left\{\overline{\sum_{j=1}^n c_j^*\Phi_{t_j};\ n\in N,\ c_j\in C^m,\ t_j\in T}\right\}.$$

定理 5.5.3 设 m 维二阶过程 z_T 有广义谱表示(5.5.2), 则

$$\mathscr{H}(z_T)^k=\left\{\int_U\tilde{F}(u)dw_u;\ \tilde{F}\in L_2(\Phi_T,\ \nu)^k\right\}.$$

$$\mathscr{R}(z_T)^k=\left\{F;\ F_t=\int_U\tilde{F}(u)d\nu_u\Phi_t^*(u),\ t\in T;\right.$$

$$\left.\tilde{F}\in L_2(\Phi_T,\ \nu)^k\right\},$$

$$S_z(F,\ G)=\int_U\tilde{F}(u)d\nu_u\tilde{G}^*(u),\ F\in\mathscr{R}(z_T)^k,\ G\in\mathscr{R}(z_T)^p,$$

$$z(F)=\int_U\tilde{F}(u)dw_u,\ F\in\mathscr{R}(z_T)^k.$$

一般有 $\mathscr{H}(z_T)\subset\mathscr{H}(w_U)$; 而 $\mathscr{H}(z_T)=\mathscr{H}(w_U)$的必要充分条件是 $L_2(\Phi_T,\ \nu)=L_2(\nu)$, 即 Φ_T 构成 $L_2(\nu)$的基.

证 由(5.5.2)及定理5.3.5之ii)和iv)，映射

$$\sum_{j=1}^{n} c_j^* \Phi_{t_j} \mapsto \sum_{j=1}^{n} c_j^* z_{t_j} = \int_U \left(\sum_{j=1}^{n} c_j^* \Phi_{t_j}(u) \right) dw_u$$

是由 Φ_T 中阵函数的所有行向量函数张成的线性空间与 $\mathscr{H}_0(z_T)$ 之间的线性保内积同构映射，因而 $\widetilde{F} \mapsto \int_U \widetilde{F}(u) dw_u$ 是 $L_2(\Phi_T, \nu)^k$ 与 $\mathscr{H}(z_T)^k$ 之间的线性保内积同构映射. 此即表明了 $x \in \mathscr{H}(z_T)^k$ 当且仅当存在(唯一的)$\widetilde{F} \in L_2(\Phi_T, \nu)^k$，使 $x = \int_U \widetilde{F}(u) dw_u$；而 $F \in \mathscr{R}(z_T)^k$ 当且仅当存在(唯一的)$\widetilde{F} \in L_2(\Phi_T, \nu)^k$，使

$$F_t = R\left(\int_U \widetilde{F}(u) dw_u,\ z_t \right) = \int_U \widetilde{F}(u) d\nu_u \Phi_t^*(u),\quad t \in T. \tag{5.5.3}$$

此时 $z(F) = \int_U \widetilde{F}(u) dw_u$，且对任一 $G \in \mathscr{R}(z_T)^p$ 有 $S_z(F, G) = R(z(F), z(G)) = \int_U \widetilde{F}(u) d\nu_u \widetilde{G}^*(u)$.

定理的最后一个断言是显然的. ∎

定理5.5.3说明了对于有广义谱表示(5.5.2)的二阶过程 z_T，其线性统计问题就相当于研究积分方程(5.5.3)对何种 F 有解及解出 $\widetilde{F} \in L_2(\Phi_T, \nu)^k$.

第 6 章

连续时间过程与线性新息定理

本章讨论连续时间随机过程，假定过程的指标集 T 为从 0 开始的有限区间和半无穷区间. 这种情形在理论上可能是更有兴趣的，但在处理方法上要比离散时间的随机序列困难得多，目前能解决的模式也狭窄得多.

记号 6.0.1 与离散时间情形的记号相对应(见记号 3.1.1)，设 $z_{[0,\infty)} \triangleq (z_t,\ t \geqslant 0)$ 为一 m 维二阶随机过程，它在 $[0,\ t]$ 上的局限为 $z_{[0,t]} = (z_s,\ 0 \leqslant s \leqslant t)$. 如果对任何 $t \in [0,\ \infty)$，$z_{[0,t]}$ 是正则过程(见定义 1.2.2)，则称 $z_{[0,\infty)}$ 为*局部正则过程*. 这时，$z_{[0,t]}$ 的生成空间 $\mathscr{H}(z_{[0,t]})$ 记作 $\mathscr{H}(z,\ t)$，或简记作 $\mathscr{H}_t$，且显然当 $s \leqslant t$ 时 $\mathscr{H}_s \subset \mathscr{H}_t$；其再生核表示空间 $\mathscr{R}(z_{[0,t]})$ 记作 $\mathscr{R}(z,\ t)$，或简记作 $\mathscr{R}_t$. 如无混淆，$[0,\ \infty]$ 上的任一矩阵值函数 $F \triangleq (F_s,\ s \geqslant 0)$ 在 $[0,\ t]$ 上的局限仍用 F 来表示，且显然当 $s \leqslant t$，$F \in \mathscr{R}_t^*$ 时，F 在 $[0,\ s]$ 上的局限属于 $\mathscr{R}_s^*$，若 $F \in \mathscr{R}_t^*$，$G \in \mathscr{R}_t^*$，它们在 $\mathscr{R}_t$ 中的内积阵记作 $S_s(F,\ G)_t$ 或简记作 $S(F,\ G)_t$，$S(F)_t \triangleq S(F,\ F)_t$，而 F 在 $\mathscr{H}_t^*$ 中的逆再生表示则记作 $z(F)_t$.

又 $t' \to t$，$t' \uparrow\uparrow t$，$t' \downarrow\downarrow t$ 分别表示 t' 趋于 t，严格上升趋于 t 和严格下降趋于 t.

§6.1 二阶连续性

定义 6.1.1 设 $z_{[0,\infty)}$ 为一 m 维局部正则过程，$t \in [0,\ \infty)$. 若有

$$\lim_{t',t''\uparrow\uparrow t} R(z_{t'}-z_{t''})=0\ (\lim_{t',t''\downarrow\downarrow t} R(z_{t'}-z_{t''})=0)$$

成立，则由 Hilbert 空间$(\mathscr{H}_t,\ R)$ 的完备性及 $z_{[0,\infty)}$的局部正则性，必有

$$z_{t'}\xrightarrow{R,\ \mathscr{L}_2} z_{t-}\in\mathscr{H}_t^m(z_{t'}\xrightarrow{R,\ \mathscr{L}_2} z_{t+}\in\bigcap_{s>t}\mathscr{H}_s^m).$$

这时称 $z_{[0,\infty)}$在 t 处有二阶左(右)极限. 进一步，若分别有 $z_{t-}=z_t$, $z_{t+}=z_t$ 或 $z_{t-}=z_t=z_{t+}$成立，则分别称 $z_{[0,\infty)}$在 t 处二阶左连续，右连续或连续. 若对一切 $t\in[0,\infty)$ 有上述性质分别成立，则分别称 $z_{[0,\infty)}$ 有二阶左(右)极限，二阶左连续，右连续或连续. 二阶连续性也称为均方连续性.

用 $\mathscr{B}_{[0,t]}$表示$[0,\ t]$上的 Borel 域，即由$[0,\ t]$中的所有开集生成的σ域. 根据定义 5.1.1，一个m维局部正则过程 $z_{[0,\infty)}$称为二阶可测的，如果对任意 $c\in\mathbf{C}^m$ 和 $t\in[0,\ \infty)$, $s|\to c^*z_s$ 是可测空间$([0,\ t],\ \mathscr{B}_{[0,t]})$映入$(\mathscr{H}_t,\ \mathscr{B}(\mathscr{H}_t))$的可测映射，亦即对任意 $x\in\mathscr{H}_t$ 和 $\alpha>0$ 有

$$\{s\in[0,\ t]: R(c^*z_s-x)<\alpha\}\in\mathscr{B}_{[0,t]}.$$

由定理 5.1.2. 这又等价于其协方差阵函数 $(R(z_t,\ z_s),\ t、s\in[0,\ \infty))$为二元 Borel 可测阵函数.

此外，若对任意 $t\in[0,\ \infty)$, Hilbert 空间 $\mathscr{H}_t$ 是可分的，即存在 $\mathscr{H}_t$ 的可数稠子集，则称 $z_{[0,\infty)}$为二阶可分过程.

引理 6.1.2 $z_{[0,\infty)}$在 t 处有二阶左(右)极限的必要充分条件是极限

$$\lim_{t',t''\uparrow\uparrow t} R(z_{t'},\ z_{t''})\ (\lim_{t',t''\downarrow\downarrow t} R(z_{t'},\ z_{t''}))$$

存在且有限，而在 t 处二阶左(右)连续的必要充分条件是上述极限等于 $R(z_t)$.

证 由引理 1.1.2 可得如下不等式：

$$\begin{aligned}&\|R(z_{t'},\ z_{s'})-R(z_t,\ z_s)\|\\&=\|R(z_{t'}-z_t,\ z_{s'}-z_s)+R(z_t,\ z_{s'}-z_s)\\&\quad+R(z_{t'}-z_t,\ z_s)\|\end{aligned}$$

$$\leqslant \sqrt{\operatorname{tr} R(z_{t'}-z_t)\operatorname{tr} R(z_{s'}-z_s)}$$
$$+\sqrt{\operatorname{tr} R(z_t)\operatorname{tr} R(z_{s'}-z_s)}$$
$$+\sqrt{\operatorname{tr} R(z_{t'}-z_t)\operatorname{tr} R(z_s)}. \qquad (6.1.1)$$

又显然有等式

$$R(z_{t'}-z_{t''})=R(z_{t'})+R(z_{t''})$$
$$-R(z_{t'},\ z_{t''})-R(z_{t''},\ z_{t'}). \qquad (6.1.2)$$

在(6.1.1)式中用 t'' 代 s' 并分别用 $(t-,\ t-)$、$(t+,\ t+)$、$(t,\ t)$ 代 $(t,\ s)$，即得必要性的证明．由(6.1.2)式则得充分性的证明(在证明二阶左(右)连续性时，用 t 代 t'')． ▌

定理 6.1.3　$z_{[0,\infty)}$ 二阶左连续的必要充分条件是：对任意 $t\in[0,\ \infty)$，有

$$\lim_{t',t''\uparrow\uparrow t} R(z_{t'},\ z_{t''})=R(z_t),$$

且此时对任何 t、$s\in[0,\ \infty)$ 有

$$\lim_{t'\uparrow\uparrow t,s'\uparrow\uparrow s} R(z_{t'},\ z_{s'})=R(z_t,\ z_s).$$

又此时 $\mathscr{R}_t^z$ 中的任一元必为 $[0,\ t]$ 上的左连续 $k\times m$ 阵函数，对于二阶右连续过程和有二阶左、右极限的过程也有相应的结果．

证　由引理 6.1.2 与(6.1.1)式，立即得到定理的前两个结论．又设 $F\in\mathscr{R}_t^z$，则 $F_s=R(z(F)_t,\ z_s)(0\leqslant s\leqslant t)$．由引理 1.1.2 即知

$$\lim_{s'\uparrow\uparrow s}\|F_{s'}-F_s\|^2=\lim_{s'\uparrow\uparrow s}\|R(z(F)_t,\ z_{s'}-z_s)\|^2$$
$$\leqslant \operatorname{tr} S(F)_t \lim_{s'\uparrow\uparrow s}\operatorname{tr} R(z_{s'}-z_s)=0.$$

即 F 在 $[0,\ t]$ 上为左连续． ▌

定理 6.1.4　设 $z_{[0,\infty)}$ 为一 m 维二阶左连续过程，则 $z_{[0,\infty)}$ 是二阶可分且可测的过程．

证　根据定理 6.1.3，此时协方差阵函数 $(R(z_t,\ z_s),\ t,s\in[0,\infty))$ 为二元左连续阵函数，所以是二元 Borel 可测阵函数．故由定理 5.1.2，$z_{[0,\infty)}$ 是二阶可测的，又由二阶左连续性，$\mathscr{H}_t=\mathscr{H}(z_{Q_t})$，其中 Q_t 表示 $[0,\ t]$ 中的有理数全体，所以 $z_{[0,\infty)}$ 又是二阶可分的．

§ 6.2　正交增量过程

定义 6.2.1　设 $w_{[0,\infty)}$ 为一 m 维局部正则过程. 若对任意的 $0\leqslant s<t<\infty$ 有 $R(w_t-w_s,\ w_s)=0$ 成立, 则称 $w_{[0,\infty)}$ 为正交增量过程. 记 $M_t\triangleq R(w_t)(t\in[0,\ \infty))$, m 阶阵函数 M 称为 $w_{[0,\infty)}$ 的核. 进一步, 设 $\mathscr{H}_t$ 由 6.0.1 给出, 若对任意 $0\leqslant s<t<\infty$, $x\in\mathscr{H}_s$ 有 $w_t\in\mathscr{H}_t^m$ 且 $R(w_t-w_s,\ x)=0$, 则称 $w_{[0,\infty)}$ 为 $\mathscr{H}$ 正交增量过程.

定理 6.2.2　设 $w_{[0,\infty)}$ 为一 m 维正交增量过程, M 为其核. 则 M 是 $[0,\ \infty)$ 上单增的 m 阶半正定 Hermite 阵函数, $w_{[0,\infty)}$ 有二阶左、右极限, 且 $w_{[0,\infty)}$ 为二阶左(右)连续的必要与充分条件是 M 为左(右)连续.

证　设 $0\leqslant s<t$, 则

$$M_t=R(w_t)=R(w_t-w_s+w_s)=R(w_t-w_s)+R(w_s)$$
$$\geqslant R(w_s)=M_s\geqslant 0.$$

所以 M 是单增的 m 阶半正定 Hermite 阵函数, 从而 M 在 $[0,\ \infty)$ 上处处有左、右极限, 因此有下式成立:

$$\lim_{t',t''\uparrow\uparrow t}\|R(w_{t'}-w_{t''})\|=\lim_{t',t''\uparrow\uparrow t}\|M_{t'}-M_{t''}\|=0,$$

$$\lim_{t',t''\downarrow\downarrow t}\|R(w_{t'}-w_{t''})\|=\lim_{t',t''\downarrow\downarrow t}\|M_{t'}-M_{t''}\|=0.$$

即 $w_{[0,\infty)}$ 有二阶左、右极限. 用同样的办法可以证明定理的最后一个结论. ▌

由这个定理知道, 若令 $v_0=w_0$, $v_t=w_{t-}(t>0)$, 则 $v_{[0,\infty)}$ 是二阶左连续的正交增量过程, 其核为 M 的左极限函数 M_-. 为方便起见, 以后我们考虑的正交增量过程总假定是二阶左连续的.

记号 6.2.3　设 $w_{[0,\infty)}$ 为一 m 维二阶左连续的正交增量过程, 从而其核 M 为左连续. 令

$$\mu([s,\ t))\triangleq M_t-M_s,\ (0\leqslant s<t<\infty).$$

由测度扩张定理, μ 可唯一扩张成可测空间 $([0,\ \infty),\ \mathscr{B}[0,\infty))$

上的m阶σ有限阵测度．又令

$$w([0,\ t))\triangleq w_t-w_0,\quad \mathscr{S}\triangleq\{[0,\ t):\ t>0\}.$$

则$\mathscr{S}$为生成$\mathscr{B}_{[0,\infty)}$的π类，$\mathscr{S}$的元都有有限μ测度，且当$n\to\infty$时，$[0,\ n)\to[0,\ \infty)$．又对任意$0\leqslant s\leqslant t<\infty$有下式成立：

$$\begin{aligned}&R(w([0,\ t)),\ w([0,\ s)))\\&=R(w_t-w_0,\ w_s-w_0)\\&=R(w_t-w_s+w_s-w_0,\ w_s-w_0)=R(w_s-w_0)\\&=M_s-M_0=\mu([0,\ s))=\mu([0,\ t)\cap[0,\ s)).\end{aligned}$$

故由定义5.3.1，$w_{\mathscr{S}}$为$\mathscr{S}$上以μ为核的m维正交随机测度．

现在对$t>0$设H是在$[0,\ t)$上对μ平方可积的$C^{k\times m}$值函数．按照测度论的习惯，记

$$\int_0^t H_s dM_s H_s^*\triangleq\int_{[0,t)} H_s\, d\mu\, H_s^*.$$

并记所有这样的H的全体为$L_2(M)_t^k$(a. e. μ相等的H视为同一个)．又记$[0,\ t]$上H对M的不定积分为$H.M$：

$$(H\cdot M)_s\triangleq\int_0^s H_r dM_r=\int_{[0,s)} H_r\, d\mu,\quad s\in[0,\ t].$$

$$\left(\frac{dM(H\cdot M)}{dM}\right)_s\triangleq H_s,\quad s\in[0,\ t).$$

由引理5.3.2知，H在$L_2(M)_t^k$中被H、M唯一确定．按照定义5.3.4与系5.3.8，可以定义$[0,\ t)$上H对$w_{\mathscr{S}}$的随机积分，并记作

$$\int_0^t H_s dw_s\triangleq\int_{[0,t)} H_s\, dw\in\mathscr{H}(w,\ t)^k.$$

对理论上严格而较广泛的讨论不感兴趣的读者，可以限于考虑被积函数H为连续的情形，而把上述随机积分理解为：对$[0,\ t]$作分割

$$0=s_0^{(n)}<s_1^{(n)}<\cdots<s_n^{(n)}=t,$$

当分割的最大步长

$$\max_{1\leqslant i\leqslant n}|s_i^{(n)}-s_{i-1}^{(n)}|\to 0$$

时，下列随机和的均方极限

$$\sum_{i=1}^{n} H_{s_i^{(n)}}(w_{s_i^{(n)}}-w_{s_{i-1}^{(n)}}) \xrightarrow{B,\ \mathscr{L}_2} \int_0^t H_s\, dw_s.$$

定理 6.2.4 设 $w_{[0,\infty)}$ 为一 m 维二阶左连续的正交增量过程，其核为 M. 则对任一 $t>0$, $w_{[0,t]}$ 的再生核表示空间 $(\mathscr{R}(w,\ t)$、$S_w(\ )_t)$ 及再生逆表示 $w(\ )_t$ 分别为

$$\mathscr{R}(w,\ t)^k \triangleq \Big\{F=F_0+\frac{dF}{dM}\cdot M;$$

$$F_0=\widetilde{F}_0M_0,\ \widetilde{F}_0\in C^{k\times m},\ \frac{dF}{dM}\in L_2(M)_t^k\Big\}$$

$$S_w(F,\ G)=F_0M_0^- G_0^*+\int_0^t\left(\frac{dF}{dM}\right)_s dM_s\left(\frac{dG}{dM}\right)_s^*,$$

$$F\in\mathscr{R}(w,\ t)^k,\ G\in\mathscr{R}(w,\ t)^l,$$

$$w(F)_t=F_0M_0^-w_0+\int_0^t\left(\frac{dF}{dM}\right)_s dw_s,\ F\in\mathscr{R}(w,\ t)^k.$$

证 注意可表

$$w_t=w_t-w_0+w_0=w([0,\ t))+w_0,$$

而最后两项是正交的. 于是由定理 1.4.1 和 5.3.3 以及 5.3.5, 即得所要的结论. ∎

例 6.2.5 设 $w_{[0,\infty)}$ 为一 m 维正交增量过程，其核 $M_t=I_mt(t\in[0,\infty))$. 例如 m 维标准 Brown 运动就具有这样的性质. 这时, $\mathscr{R}(w,\ t)^k$ 由 $[0,\ t]$ 上满足下述条件的 $C^{k\times m}$ 值函数 F 全体组成: F 的每一元都是绝对连续函数, 且其导数(几乎处处存在)在 $[0,\ t]$ 上对 Lebesgue 测度为平方可积. 如果用 $\dot{F}$ 表示对 F 的每一元求导所得的阵函数, 则有

$$\int_0^t \operatorname{tr}(\dot{F}_s\dot{F}_s^*)ds<\infty,$$

且
$$S_w(F)_t=\int_0^t \dot{F}_s\dot{F}_s^* ds,\ w(F)_t=\int_0^t \dot{F}_s\, dw_s.$$

正交增量过程也可以定义在整个实轴 $(-\infty,\ \infty)$ 上或任意有穷或无穷区间上, 并有与 6.2.4 类似的定理成立. 根据定义 5.5.2, 如果某一二阶过程 z_T 对某个正交增量过程有广义谱表示(例如二阶平稳过程), 则可由上述定理和定理 5.5.3 定出 z_T 的再

生核表示空间及其逆再生表示. J. Hajek[1] 曾举了一些有趣且有用的例子,因为这些内容不是本书的重点,我们在此不加列举,感兴趣的读者可参见参考文献[1].

定理 6.2.6 设 $z_{[0,\infty)}$为一 m 维局部正则过程, F 为$[0,\infty)$上的 $C^{k\times m}$ 值函数. 若对任意的 $t\in[0,\infty)$, $F\in\mathscr{R}_t^z$, 则 $z(F)_{[0,\infty)}\triangleq(z(F)_t,\ t\in[0,\infty))$为 k 维 $\mathscr{H}$ 正交增量过程, 其核为 $S(F)\triangleq(S(F)_t,\ t\in[0,\infty))$. 又若 $z_{[0,\infty)}$为二阶左连续, 则 $z(F)_{[0,\infty)}$也为二阶左连续.

证 由定义, $z(F)_t\in\mathscr{H}_t^k$. 又对任意的 $0\leqslant r\leqslant s<t$, 有

$$R(z(F)_t-z(F)_s,\ z_r)=F_r-F_r=0.$$

因此对任意的 $x\in\mathscr{H}_s$ 有 $R(z(F)_t-z(F)_s,\ x)=0$. 由定义 6.2.1, $z(F)_{[0,\infty)}$为 $\mathscr{H}$ 正交增量过程,其核为 $R(z(F)_t)=S(F)_t$, $t\in[0,\infty)$. 由定理 6.2.2, 二阶左、右极限 $z(F)_{t-}$、$z(F)_{t+}$ 都存在, $z(F)_{t-}\in\mathscr{H}_t^k$, 且对任意的 $s<t$,

$$R(z(F)_{t-},\ z_s)=\lim_{t'\uparrow\uparrow t}R(z(F)_{t'},\ z_s)=F_s.$$

又若 $z_{[0,\infty)}$为二阶左连续,则由定理 6.1.3 更有

$$R(z(F)_{t-},\ z_t)=\lim_{t'\uparrow\uparrow t}R(z(F)_{t-},\ z_{t'})=\lim_{t'\uparrow\uparrow t}F_{t'}=F_t.$$

由再生逆表示的唯一性即知 $z(F)_{t-}=z(F)_t$, 即 $z(F)_{[0,\infty)}$为二阶左连续. ∎

为了下一节的需要,给出下面的一个定理.

定理 6.2.7 设 $z_{[0,\infty)}$与 $x_{[0,\infty)}$分别为m维与 n 维局部正则过程,令 $y_t\triangleq z(Z^{x_t})_t$ $(t\in[0,\infty))$. 若 $z_{[0,\infty)}$ 为二阶左连续而 $x_{[0,\infty)}$ 为二阶可分可测,则 $y_{[0,\infty)}$ 为 n 维二阶可测过程. 特别是若 $x_{[0,\infty)}$ 为二阶左连续, 则 $y_{[0,\infty)}$ 也为二阶左连续; $x_{[0,\infty)}$ 有二阶右极限,则 $y_{[0,\infty)}$ 也有二阶右极限.

证 先设 $x_{[0,\infty)}=x1_A$, 其中 $x\in\mathscr{L}_2^n$, $A\in\mathscr{B}_{[0,\infty)}$; 即设

$$x_t=\begin{cases}x, & t\in A;\\ 0, & t\in[0,\infty)\setminus A.\end{cases}$$

这时显然有 $y_{[0,\infty)}=z(Z^x)_A=z(Z^x)_{[0,\infty)}1_A$. 由假设和定理 6.2.6,

$z(Z^x)_{[0,\infty)}$为二阶左连续，因而由定理 6.1.4，为二阶可测的. 所以 $y_{[0,\infty)}$是二阶可测过程.

对于一般情形，任意固定 $t>0$, $c\in C^n$. 由 $x_{[0,\infty)}$的二阶可分可测性假设(见定义 6.1.1)，设可数集$\{u_k\}$稠于 $\mathscr{H}(x, t)$. 对任意 k、$j\in N$，令

$$B_{jk}\triangleq\{s\in[0,\ t]: R(c^*x_s-u_k)<2^{-j}\},$$

$$B_{jk}^c\triangleq[0,\ t]\setminus B_{jk},$$

$$A_{j1}\triangleq B_{j1},\ A_{jk}\triangleq B_{jk}B_{jk-1}^c\cdots B_{j1}^c,\ (k>1).$$

则
$$A_{jk},\ B_{jk}\in\mathscr{B}_{[0,t]},\ A_{jk}\subset B_{jk},$$

$$[0,\ t]=\bigcup_k B_{jk}=\sum_k A_{jk},\quad (j\in N).$$

又令
$$u_{[0,t]}^{(j)}\triangleq\sum_k u_k\mathbf{1}_{A_{jk}},\ v_{[0,t]}^{(j)}\triangleq\sum_k z(Z^{u_k})_{A_{jk}}.$$

则由上知 $u_{[0,t]}^{(j)}$ 与 $v_{[0,t]}^{(j)}$ 都是$[0,\ t]$上的 1 维二阶可测过程. 对任一 $s\in[0,\ t]$, $j\in N$, 必存在 $k(j)\in N$, 使 $s\in A_{jk(j)}\subset B_{jk(j)}$, 故当 $j\to\infty$ 时，

$$\begin{aligned}R(c^*y_s-v_s^{(j)})&=R(z(Z^{c^*x_t})_s-z(Z^{u_{k(j)}})_s)\\&=R(z(Z^{c^*x_t-u_{k(j)}})_s)\leqslant R(c^*x_s-u_{k(j)})\\&<2^{-j}\to 0.\end{aligned}$$

即二阶过程列 $v_{[0,t]}^{(j)}$ 处处均方收敛于 $c^*y_{[0,t]}$. 由 c 和 t 的任意性，即证得 $y_{[0,\infty)}$为二阶可测过程.

若 $x_{[0,\infty)}$为二阶左连续，则由定理 6.2.6,

$$\begin{aligned}&\lim_{t'\uparrow\uparrow t}R(y_t-y_{t'})\\&\quad=\lim_{t'\uparrow\uparrow t}R(z(Z^{x_t})_t-z(Z^{x_t})_{t'}+z(Z^{x_t-x_{t'}})_{t'})\\&\quad\leqslant\lim_{t'\uparrow\uparrow t}(S(Z^{x_t})_t-S(Z^{x_t})_{t'})\\&\qquad+\lim_{t'\uparrow\uparrow t}R(x_t-x_{t'})=0.\end{aligned}$$

故 $y_{[0,\infty)}$也为二阶连续. 若 x_{t+}存在，记

$$z(Z^{x_{t+}})_{t+}\triangleq\lim_{t''\downarrow\downarrow t}z(Z^{x_{t+}})_{t''}$$

(由定理 6.2.6 和 6.2.2，此均方极限必然存在). 则依同理可证

$$\lim_{s'' \downarrow \downarrow t} R(y_{t''} - z(Z^{x(\cdot)})_{t+}) = 0.$$

即由 $x_{[0,\infty)}$ 有二阶右极限可推出 $y_{[0,\infty)}$ 也有二阶右极限. ∎

对 $y_{[0,\infty)}$ 的二阶可测性的较繁琐证明不感兴趣的读者只须理解本定理第二部分的结论即可(因为左连续性蕴含可测性).

§6.3 线性新息定理

上一节的定理 6.2.4 给出了正交增量过程的再生核表示空间及其逆再生表示的完整解答. 正交增量过程在连续时间过程类中所起的作用正如同正交随机序列在离散时间过程(即随机序列)类中所起的作用一样. 在离散时间的情形, 我们曾对局部正则序列定义了它的新息序列(见定义 3.1.2), 并且证明了后者是正交随机序列(见引理 3.1.3 及其注), 从而把前者的再生核表示的解答归结为后者(见整个第 3 章). 这个方法实际上就是 Hilbert 空间中元的序列的 Schmidt 正交化方法. 但是对于连续时间的随机过程, 一般却并不存在这样的"正交化", 只是对一类特殊而很有用的过程, 才有所谓"线性新息定理"成立(见后而定理 6.3.3), 它的实质就是使一个连续时间过程与一个正交增量过程"线性等效". 就作者所知, 它的第一个严格证明是由原苏联学者 Y.A.Rozanov[7] 给出的, 几乎用了一本小册子的篇幅, 工具也相当复杂. 本节我们将给出它的一个相当简洁的证明, 而且还推广到高维情形并减弱了原来定理的条件.

定义 6.3.1 设 $z_{[0,\infty)}$ 为一 m 维局部正则过程, 若存在一 m 维正交增量过程 $\varepsilon_{[0,\infty)}$ 使对一切 $t \in [0, \infty)$ 有 $\mathscr{H}_t \triangleq \mathscr{H}(z, t) = \mathscr{H}(\varepsilon, t)$, 则称 $\varepsilon_{[0,\infty)}$ 为 $z_{[0,\infty)}$ 的(线性)新息过程.

为了证明线性新息定理, 需要一个线性积分方程解的结果.

引理 6.3.2 设 M 为 $[0, \infty)$ 上单调增、左连续的 m 阶 Hermite 阵函数. $K \triangleq (K(t, s): 0 \leqslant s < t < \infty)$ 是其定义域上二元 Borel 可测的 m 阶阵函数, 且对任意 $t \in [0, \infty)$ 满足

$$\alpha_t \triangleq \mathrm{tr}\left[\int_0^t \left(\int_0^s K(s, r) dM_r K^*(s, r)\right) dM_s\right] < \infty.$$

又用 $\mathscr{X}_M^{m\times k}$ 表示 $[0, \infty)$ 上这样的 Borel 可测 $C^{m\times k}$ 值函数 X 的全体：对任意的 $t\in[0, \infty)$，$X^*\in L_2(M)_t^k$ (a. e. M 相等的 X 视为相同)．

若 $V\in\mathscr{X}_M^{m\times k}$，则如下的 Volterra 型积分方程

$$X_t = V_t - \int_0^t K(t, s) dM_s X_s \quad (t\in[0, \infty)) \tag{6.3.1}$$

在 $\mathscr{X}_M^{m\times k}$ 中有解．进一步，若 M 连续，则解还是唯一的．

证 我们逐步构造一个 $C^{m\times k}$ 值函数列 $\{X^{(n)}\}$ 来逼近 (6.3.1) 的解，为此，令

$$X_t^{(0)} \triangleq V_t, \quad X_t^{(n)} \triangleq V_t - \int_0^t K(t, s) dM_s X_s^{(n-1)}$$
$$(t\in[0, \infty),\ n\geqslant 1).$$

下面用归纳法证明：

$X^{(n)}\in\mathscr{X}_M^{m\times k}$，且对任意的 $t\in[0, \infty)$，有下式成立：

$$\beta_t^{(n)} \triangleq \mathrm{tr}\left[\int_0^t (X_s^{(n)} - X_s^{(n-1)})^* dM_s (X_s^{(n)} - X_s^{(n-1)})\right]$$
$$\leqslant \frac{\alpha_t^n}{n!}\mathrm{tr}\left(\int_0^t V_s^* dM_s V_s\right). \tag{6.3.2}$$

首先，若记 $X^{(-1)}\triangleq 0$，则上述结论对 $n=0$ 显然成立．其次，设对 n 上述结论已成立，则由引理 5.1.6 和归纳假设即得

$$\beta_t^{(n+1)} = \mathrm{tr}\left[\int_0^t \left(\int_0^s K(s, r) dM_r (X_r^{(n)} - X_r^{(n-1)})\right)^*\right.$$
$$\left.\times dM_s \left(\int_0^s K(s, r) dM_r (X_r^{(n)} - X_r^{(n-1)})\right)\right]$$
$$= \int_0^t \left|\int_0^s \left(\frac{dM}{d\,\mathrm{tr}M}\right)_s^{\frac{1}{2}} K(s, r) dM_r (X_r^{(n)} - X_r^{(n-1)})\right|^2 d\,\mathrm{tr}M_s$$
$$\leqslant \int_0^t \left\{\mathrm{tr}\left[\int_0^s \left(\frac{dM}{d\,\mathrm{tr}M}\right)_s^{\frac{1}{2}} K(s, r) dM_r K^*(s, r) \left(\frac{dM}{d\,\mathrm{tr}M}\right)_s^{\frac{1}{2}}\right]\right.$$
$$\left.\times\mathrm{tr}\left[\int_0^s (X_r^{(n)} - X_r^{(n-1)}) dM_r (X_r^{(n)} - X_r^{(n-1)})\right]\right\} d\,\mathrm{tr}M_s$$

$$=\int_0^t \beta_s^{(n)} d\alpha_s \leqslant \int_0^t \frac{\alpha_s^n}{n!} \operatorname{tr}\left(\int_0^s V_r^* dM_r V_r\right) d\alpha_s$$

$$\leqslant \frac{\alpha_t^{n+1}}{(n+1)!} \operatorname{tr}\left(\int_0^t V_r^* dM_r V_r\right)$$

(注意 α 为 $[0, \infty)$ 上的非负左连续增函数).

这就证明了(6.3.2), 且由此可从 $X^{(n-1)} \in \mathscr{X}_M^{m\times k}$ 推出 $X^{(n)} \in \mathscr{X}_M^{m\times k}$. 又注意由著名的 Stirling 公式可知

$$n! > n^n e^{-n} \sqrt{2\pi},$$

从而

$$\sqrt{\frac{\alpha_t^n}{n!}} < (2\pi)^{-\frac{1}{4}} \left(\frac{e\alpha_t}{n}\right)^{\frac{n}{2}}.$$

因此由(6.3.2)显然有

$$\sum_{n=0}^{\infty} \sqrt{\beta_t^{(n)}} \leqslant \sum_{n=0}^{\infty} \sqrt{\frac{\alpha_t^n}{n!}} \sqrt{\operatorname{tr}\left(\int_0^t V_s^* dM_s V_s\right)} < \infty.$$

于是由引理 5.1.6 的(5.1.2)式, 对任意的 $t \in [0, \infty)$, 当 $n' > n \to \infty$ 时, 有

$$\left(\operatorname{tr}\left[\int_0^t (X_s^{(n')} - X_s^{(n)})^* dM_s (X_s^{(n')} - X_s^{(n)})\right]\right)^{\frac{1}{2}}$$

$$= \left(\operatorname{tr}\left[\int_0^t \left(\sum_{j=n+1}^{n'} (X_s^{(j)} - X_s^{(j-1)})\right)^* dM_s \right.\right.$$

$$\left.\left. \times \left(\sum_{j=n+1}^{n'} (X_s^{(j)} - X_s^{(j-1)})\right)\right]\right)^{\frac{1}{2}}$$

$$\leqslant \sum_{j=n+1}^{n'} \sqrt{\beta_t^{(j)}} \to 0.$$

由此以及 $L_2(M)_t^k$ 在范数

$$\|H\|_{Mt} \triangleq \left(\operatorname{tr}\left(\int_0^t H_s dM_s H_s^*\right)\right)^{\frac{1}{2}}$$

之下的完备性(易从普通 L_2 空间的完备性推知), 必存在 $[0, t)$ 上唯一的 $C^{m\times k}$ 值函数 X, 使 $X^* \in L_2(M)_t^m$, 且当 $n \to \infty$ 时,

$$\operatorname{tr}\left[\int_0^t (X_s - X_s^{(n)})^* dM_s (X_s - X_s^{(n)})\right] \to 0.$$

由 t 的任意性和极限函数的唯一性知, 上述 X 可在 $[0, \infty)$ 上唯一确定, 且 $X \in \mathscr{X}_M^{m\times k}$.

为证这样确定的 X 必满足方程(6.3.1)，暂时记

$$Y_t=V_t-\int_0^t K(t,\ s)dM_sX_s\quad(t\in[0,\ \infty)).$$

仿(6.3.2)式的证明可得：当 $n\to\infty$ 时，

$$\operatorname{tr}\left[\int_0^t (Y_s-X_s^{(n)})^* dM_s(Y_s-X_s^{(n)})\right]$$

$$\leqslant \alpha_t \operatorname{tr}\left[\int_0^t (X_s-X_s^{(n-1)})^* dM_s(X_s-X_s^{(n-1)})\right]\to\infty.$$

由极限函数的唯一性即知 $Y=X$，亦即 X 为方程(6.3.1)的解.

现在设 M 连续，则 α 是$[0,\ \infty)$上的非负连续增函数．再设 X、$X'\in\mathscr{X}_M^{m\times k}$，且都是(6.3.1)的解. 记

$$0\leqslant\beta_t\triangleq\operatorname{tr}\left[\int_0^t (X_s-X_s')^* dM_s(X_s-X_s')\right],$$

则仿(6.3.2)式的证明可得不等式

$$\beta_t\leqslant\int_0^t\beta_s d\alpha_s.$$

把上式右端积分号下的 β 再用此不等式代入并交换积分次序，这样反复迭代下去，即可归纳地得到：

$$0\leqslant\beta_t\leqslant\int_0^t\beta_s d\alpha_s\leqslant\int_0^t\left(\int_0^s\beta_r d\alpha_r\right)d\alpha_s$$

$$=\int_0^t\left(\int_r^t d\alpha_s\right)\beta_r d\alpha_r=\int_0^t(\alpha_t-\alpha_r)\beta_r d\alpha_r$$

$$\leqslant\int_0^t(\alpha_t-\alpha_r)\left(\int_0^r\beta_s d\alpha_s\right)d\alpha_r=\int_0^t\left(\int_s^t(\alpha_t-\alpha_r)d\alpha_r\right)\beta_s d\alpha_s$$

$$=\int_0^t\frac{(\alpha_t-\alpha_s)^2}{2}\beta_s d\alpha_s\leqslant\cdots\leqslant\int_0^t\frac{(\alpha_t-\alpha_s)^n}{n!}\beta_s d\alpha_s.$$

令 $n\to\infty$，即得 $\beta_t=0\ (t\in[0,\ \infty))$，亦即 $X=X'$. 这就证明了解的唯一性. ▌

注：值得注意的是：上述迭代方法只有当 α 为连续增函数时才是正确的. 事实上，若 M 不连续，则方程(6.3.1)的解未必唯一. 举一个简单的例子：设 $m=k=1$，在(6.3.1)中置 $V=0$，$K=-1$，

$$M_t=\begin{cases}0,& 0\leqslant t\leqslant 1,\\ 1,& 1<t<\infty;\end{cases}\qquad X_t=\begin{cases}0,& 0\leqslant t\leqslant 1,\\ a,& 1<t<\infty.\end{cases}$$

其中 a 是任何数. 显然所有这样的 X 都满足(6.3.1), 而对不同的 a, X 并非 a.e. M 相等.

下面转向叙述和证明线性新息定理. 设 $w_{[0,\infty)}$ 为一 m 维二阶左连续的正交增量过程,其核为 M, $x_{[0,\infty)}$为一 m 维二阶可分可测的局部正则过程,且对任意 $t\in[0, \infty)$,

$$\operatorname{tr}\left(\int_0^t R(x_s)dM_s\right)<\infty.$$

这样的 $x_{[0,\infty)}$ 称为对 M 二阶局部可积(注意: 若 $x_{[0,\infty)}$ 为二阶左连续且有二阶右极限的局部正则过程, 则由定理 6.1.3 和 6.1.4, 上述条件自然成立). 又设对任意 $0\leqslant r\leqslant s<t$, $R(w_t-w_s,\ x_r)=0$, 这意味着 $w_{[0,\infty)}$的增量与 $x_{[0,\infty)}$的"线性历史"不相关(即正交). 在上述诸假设之下,考虑如下形式的 m 维过程 $z_{[0,\infty)}$:

$$z_t=\int_0^t dM_s x_s+w_t \quad (t\in[0,\ \infty)). \tag{6.3.3}$$

根据假设和定义 5.2.3, 上式右边的随机积分是有意义的. 这种形式的随机过程在应用上很常见,它的第一项可理解为对 $x_{[0,t)}$的某种积累(积分),而 w_t 则是量测噪声. 又记

$$\Delta M_t\triangleq M_{t+}-M_t \quad (t\in[0,\ \infty)),$$

由单调函数的性质, m 阶 Hermite 阵 $\Delta M_t\geqslant 0$, 且最多只在可数个 t 处不为 0.

下面就是 $z_{[0,\infty)}$的线性新息定理:

定理 6.3.3 设 $z_{[0,\infty)}$由(6.3.3)式给出,且满足前述假设. 对任意 $t\in[0, \infty)$,令

$$y_t\triangleq z(Z^{x_t})_t,$$

$$P_t\triangleq R(x_t-y_t)=R(x_t)-S(Z^{x_t})_t,$$

$$\varepsilon_t\triangleq z_t-\int_0^t dM_s y_s, \tag{6.3.4}$$

$$N_t\triangleq M_t+\sum_{0\leqslant s<t}\Delta M_s P_s \Delta M_s=\int_0^t (I+\Delta M_s P_s)dM_s. \tag{6.3.5}$$

则 $\varepsilon_0=z_0=w_0$, $\varepsilon_{[0,\infty)}$为以 N 为核的二阶左连续 $\mathscr{H}$ 正交增量过

程，并且是 $z_{[0,\infty)}$ 的(线性)新息过程. 此时 $z_{[0,t)}$ 的再生核表示空间 $(\mathscr{R}_t, S(\ )_t)$ 及其再生逆表示 $z(\ )_t$ 分别为

$$\mathscr{R}_t^k=\left\{F=F_0+\frac{dF}{dN}\cdot N: F_0=\bar{F}_0M_0,\ \bar{F}_0\in C^{k\times m},\ \frac{dF}{dN}\in L_2(N)_t^k\right\}, \tag{6.3.6}$$

$$S(F,\ G)_t=F_0M_0^-G_0^*+\int_0^t\left[\left(\frac{dF}{dN}\right)_s-X_s^{F*}\right]dN_s\left[\left(\frac{dG}{dN}\right)_s-X_s^{G*}\right]^*,\quad F\in\mathscr{R}_t^k,\ G\in\mathscr{R}_t^l. \tag{6.3.7}$$

$$z(F)_t=F_0M_0^-z_0+\int_0^t\left[\left(\frac{dF}{dN}\right)_s-X_s^{F*}\right]d\varepsilon_s. \tag{6.3.8}$$

其中 X_s^F, $s\in[0,\ t)$ 为如下 Volterra 型积分方程的唯一解:

$$X_s^F=V_s^F-\int_0^t K(s,\ r)dN_rX_r^F, \tag{6.3.9}$$

这里

$$K(t,\ s)\triangleq(I+P_t\Delta M_t)^{-1}\times\left[\left(\frac{dZ^{x_t}}{dN}\right)_s-S(Z^{x_t},\ Z^{x_t})_s(I+\Delta M_sP_s)^{-1}\right]\quad(0\leqslant s<t<\infty), \tag{6.3.10}$$

$$V_s^F\triangleq(I+P_s\Delta M_s)^{-1}R(x_s,\ z_0)M_0^-F_0^*+\int_0^s K(s,\ r)dN_r\left(\frac{dF}{dN}\right)_r^*\quad(s\in[0,\ t)). \tag{6.3.11}$$

证　首先，由定理 6.2.7 知 $y_{[0,\infty)}$ 为二阶可测，且有

$$\begin{aligned}\operatorname{tr}\left(\int_0^t R(y_s)dM_s\right)&=\int_0^t\operatorname{tr}\left[\left(\frac{dM}{d\operatorname{tr}M}\right)_s^{\frac12}R(y_s)\left(\frac{dM}{d\operatorname{tr}M}\right)_s^{\frac12}\right]d\operatorname{tr}M_s\\&\leqslant\int_0^t\operatorname{tr}\left[\left(\frac{dM}{d\operatorname{tr}M}\right)_s^{\frac12}R(x_s)\left(\frac{dM}{d\operatorname{tr}M}\right)_s^{\frac12}\right]d\operatorname{tr}M_s\\&=\operatorname{tr}\left(\int_0^t R(x_s)dM_s\right)<\infty.\end{aligned}$$

所以 $y_{[0,\infty)}$ 对 M 二阶局部可积，从而(6.3.4)式中的随机积分是有意义的. 显然 $\varepsilon_0=x_0-w_0$, 且对任意 $t\in[0,\infty)$ 有 $\mathscr{H}(\varepsilon,t)\subset\mathscr{H}_t$.

记 $\tilde{x}_t\triangleq x_t-y_t$, 将(6.3.3)式代入(6.3.4)式，得到

$$\varepsilon_t=w_t+\int_0^t dM_s\tilde{x}_s. \tag{6.3.12}$$

于是由定理 5.2.4 中 ii)和假设条件，对任意的 $0\leqslant r\leqslant s<t$ 有下式成立：

$$\begin{aligned}
&R(\varepsilon_t-\varepsilon_s,\ z_r)\\
&=R(w_t-w_s,\ z_r)+R\left(\int_s^t dM_\tau\tilde{x}_\tau,\ z_r\right)\\
&=R(w_t-w_s,\ w_r)+\int_0^r R(w_t-w_s,\ x_\tau)dM_\tau\\
&\quad+\int_s^t dM_\tau R(x_\tau-y_\tau,\ z_r)=0.
\end{aligned}$$

这就证明了 $\varepsilon_{[0,\infty)}$ 是 $\mathscr{H}$ 正交增量过程(见定义 6.2.1). 至于其二阶左连续性，则可由 $w_{[0,\infty)}$ 的同样性质以及定理 5.2.4 的 V)，当 $t'\uparrow\uparrow t$ 时，

$$\int_0^{t'}dM_sy_s=\int_0^t I_{[0,t')}(s)dM_sy_s\xrightarrow{\mathscr{L}_2}\int_0^t dM_sy_s$$

推出. 为计算 $\varepsilon_{[0,\infty)}$ 的核，利用(6.3.12)式、定理 5.2.4 的 ii)以及假设条件，

$$\begin{aligned}
R(\varepsilon_t)&=R(w_t)+\int_0^t dM_sR(\tilde{x}_s,\ w_t-w_s+w_s)\\
&\quad+\int_0^t R(w_t-w_r+w_r,\ \tilde{x}_r)dM_r\\
&\quad+\int_0^t dM_s\left(\int_0^t R(\tilde{x}_s,\ \tilde{x}_r)dM_r\right)\\
&=M_t+\int_0^t dM_s\ R(\tilde{x}_s,\ w_s)+\int_0^t R(w_r,\ \tilde{x}_r)dM_r\\
&\quad+\int_0^t dM_s\left(\int_0^t R(\tilde{x}_s,\ \tilde{x}_r)dM_r\right)\\
&\triangleq M_t+J_1+J_2+J_3,
\end{aligned}$$

其中(计算 J_2 的最后一式时交换积分次序)

$$J_1=\int_0^t dM_s R\left(\tilde{x}_s,\ \varepsilon_s-\int_0^s dM_r\tilde{x}_r\right)$$

$$=-\int_0^t dM_s\left(\int_{[0,s)} R(\tilde{x}_s,\ \tilde{x}_r)\right)dM_r,$$

$$J_2=-\int_0^t\left(\int_{[0,r)} dM_s R(\tilde{x}_s,\ \tilde{x}_r)\right)dM_r$$

$$=-\int_0^t dM_s\left(\int_{(s,t)} R(\tilde{x}_s,\ \tilde{x}_r)dM_r\right).$$

由此得

$$J_1+J_2+J_3=\int_0^t dM_s(R(\tilde{x}_s)\Delta M_s)=\int_0^t dM_s(P_s\Delta M_s)$$

$$=\sum_{0\leqslant s<t}\Delta M_s P_s\Delta M_s=\int_0^t(\Delta M_s P_s)dM_s.$$

这就证明了 $\varepsilon_{[0,\infty)}$的核为

$$R(\varepsilon_t)=M_t+\sum_{0\leqslant s<t}\Delta M_s P_s\Delta M_s\triangleq N_t\ \ (t\in[0,\ \infty)).$$

其次，任意固定 t，设 $F\in\mathscr{R}_t^k$，则

$$F_0\in\mathscr{R}_0^k\triangleq\mathscr{R}(z_0)^k=\mathscr{R}(\varepsilon_0)^k.$$

由定理 1.4.1，必存在 $\bar{F}_0\in C^{k\times m}$，使 $F_0=\bar{F}_0M_0$. 由定理 2.1.3，必存在唯一的 $u\in\mathscr{H}(\varepsilon,\ t)^k$，使

$$R(z(F)_t,\ \varepsilon_s)=R(u,\ \varepsilon_s)\ \ (0\leqslant s\leqslant t).$$

又由定理 6.2.4，必存在 $H\in L_2(N)_t^k$，使

$$u=F_0M_0^-z_0+\int_0^t H_s\,d\varepsilon_s.$$

于是由(6.3.4)式和定理 5.3.5 的 ii)，对任意 $s\in[0,\ t]$，有

$$F_s=R(z(F)_t,\ z_s)=R\left(z(F)_t,\ \varepsilon_s+\int_0^s dM_r y_r\right)$$

$$=R(u,\ \varepsilon_s)+\int_0^s R(z(F)_t-z(F)_r+z(F)_r,\ y_r)dM_r$$

$$=R\left(F_0M_0^-\varepsilon_0+\int_0^t H_r\,d\varepsilon_r,\ \varepsilon_0+\int_0^s d\varepsilon_r\right)$$

$$+\int_0^s R(z(F)_r,\ z(Z^{x_r})_r)dM_r$$

$$=F_0+\int_0^s H_r\,dN_r+\int_0^s S(F,\ Z^{x_r})_r dM_r.$$

注意由(6.3.5)式,

$$N_t=\int_0^t(I+\Delta M_sP_s)dM_s, \quad (t\in[0,\ \infty)).$$

其中 ΔM_s 和 P_s 都是半正定 Hermite 阵. 由熟知的矩阵与行列式的性质:

$$\det(I+\Delta M_sP_s)=\det(I+P_s^{\frac{1}{2}}\Delta M_sP_s^{\frac{1}{2}})\geqslant 1,$$

故 $(I+\Delta M_sP_s)$ 为可逆阵. 因此又可表

$$F_s=F_0+\int_0^s[H_r+S(F,\ Z^{*r})_r(I+\Delta M_rP_r)^{-1}]dN_r$$

$$\triangleq F_0+\int_0^s\left(\frac{dF}{dN}\right)_r dN_r \quad (s\in[0,\ t]). \tag{6.3.13}$$

下面证明: $\frac{dF}{dN}\in L_2(N)_t^*$. 为此,注意由矩阵不等式

$$\begin{aligned}&(I+\Delta M_sP_s)\Delta M_s(I+P_s\Delta M_s)\\&\quad=\Delta M_s+2\Delta M_sP_s\Delta M_s+\Delta M_sP_s\Delta M_sP_s\Delta M_s\\&\quad\geqslant\Delta M_s+\Delta M_sP_s\Delta M_s\end{aligned}$$

可得

$$\begin{aligned}\Delta M_s\geqslant&(I+\Delta M_sP_s)^{-1}\\&\times(\Delta M_s+\Delta M_sP_s\Delta M_s)(I+P_s\Delta M_s)^{-1}.\end{aligned}$$

又用 $M_t^c\triangleq M_t-\sum_{0\leqslant s<t}\Delta M_s$ 表 M 的"连续部分", 则

$$N_t^c=N_t-\sum_{0\leqslant s<t}\Delta N_s=N_t-\sum_{0\leqslant s<t}(\Delta M_s+\Delta M_sP_s\Delta M_s)=M_t^c.$$

于是有

$$\begin{aligned}&\int_0^t(I+\Delta M_sP_s)^{-1}dN_s(I+P_s\Delta M_s)^{-1}\\&\quad=N_t^c+\sum_{0\leqslant s<t}(I+\Delta M_sP_s)^{-1}\\&\qquad\times(\Delta M_s+\Delta M_sP_s\Delta M_s)(I+P_s\Delta M_s)^{-1}\\&\quad\leqslant M_t^c+\sum_{0\leqslant s<t}\Delta M_s=M_t \quad (t\in[0,\ \infty)).\end{aligned} \tag{6.3.14}$$

由此及引理 1.1.2 即知

$$\mathrm{tr}\left[\int_0^t S(F,Z^{*s})_s(I+\Delta M_sP_s)^{-1}dN_s(I+P_s\Delta M_s)^{-1}S(Z^{*s},F)_s\right]$$

$$\leqslant \operatorname{tr}\left[\int_0^t S(F, Z^{x_s})_s dM_s S(Z^{x_s}, F)_s\right]$$

$$= \operatorname{tr}\left[\int_0^t R(z(F)_t, y_s) dM_s R(y_s, z(F)_t)\right]$$

$$= \int_0^t \left\| R\left(z(F)_t, \left(\frac{dM}{d\operatorname{tr}M}\right)_s^{\frac{1}{2}} y_s\right)\right\|^2 d\operatorname{tr}M_s$$

$$\leqslant \operatorname{tr}S(F)_t \operatorname{tr}\left(\int_0^t R(y_s) dM_s\right) < \infty.$$

将此结果用于(6.3.13)式中 $\frac{dF}{dN}$ 的表达式，并且注意已知 $H\in L_2(N)_t^k$，即推得 $\frac{dF}{dN}\in L_2(N)_t^k$. 换言之，我们已经证得(6.3.6)式的一半：

$$\mathscr{R}_t^k \subset \left\{ F = F_0 + \frac{dF}{dN}\cdot N: F_0 = \bar{F}_0 M_0,\right.$$

$$\left.\bar{F}_0 \in C^{k\times m},\ \frac{dF}{dN}\in L_2(N)_t^k\right\}.$$

再次，用上段同样的思路，分别由定理 2.1.3 与定理 6.2.4，对任意 $t\in[0, \infty)$，必存在 $v_t\in\mathscr{H}(\varepsilon, t)^m$，使

$$R(y_t, \varepsilon_s) = R(x_t, \varepsilon_s) = R(v_t, \varepsilon_s),\ (0\leqslant s\leqslant t), \tag{6.3.15}$$

且存在 $H(t, s)$，$0\leqslant s<t<\infty$，使 $H(t, \cdot)\in L_2(N)_t^m$，而

$$v_t = R(x_t, z_0)M_0^- z_0 + \int_0^t H(t, s) d\varepsilon_s \tag{6.3.16}$$

依照(6.3.13)及其前式同样的推理得到，对任意 $s\in[0, t]$，

$$Z_s^{x_t} = R(y_t, z_s)$$

$$= R(x_t, z_0) + \int_0^s H(t, r) dN_r + \int_0^s S(Z^{x_t}, Z^{x_r})_r dM_r$$

$$= R(x_t, z_0)$$

$$+ \int_0^s \left[H(t, r) + S(Z^{x_t}, Z^{x_r})(I + \Delta M_r P_r)^{-1}\right] dN_r$$

即

$$\left(\frac{dZ^{x_t}}{dN}\right)_s = H(t,s)+S(Z^{x_t},\ Z^{x_s})(I+\Delta M_s P_s)^{-1},$$

$$(0\leqslant s<t<\infty).$$

由(6.3.10)式，令

$$K(t,\ s)=(I+P_t\Delta M_t)^{-1}H(t,\ s),$$

下面验证它满足引理 6.3.2 对 K 的假设条件：事实上，由(6.3.16)、(6.3.14)和(6.3.15)式，

$$\begin{aligned}
&\operatorname{tr}\left[\int_0^t\left(\int_0^s K(s,\ r)dN_r K^*(s,\ r)\right)dN_s\right]\\
&=\operatorname{tr}\left[\int_0^t (I+P_s\Delta M_s)^{-1}\left(\int_0^s H(s,\ r)dN_r H^*(s,r)\right)\right.\\
&\quad\left.\times(I+\Delta M_s P_s)^{-1}dN_s\right]\\
&\leqslant\operatorname{tr}\left[\int_0^t (I+P_s\Delta M_s)^{-1}R(v_s)(I+\Delta M_s P_s)^{-1}dN_s\right]\\
&=\operatorname{tr}\left[\int_0^t R(v_s)^{\frac{1}{2}}(I+\Delta M_s P_s)^{-1}dN_s(I+P_s\Delta M_s)^{-1}R(v_s)^{\frac{1}{2}}\right]\\
&\leqslant\operatorname{tr}\left(\int_0^t R(v_s)dM_s\right)\leqslant\operatorname{tr}\left(\int_0^t R(x_s)dM_s\right)<\infty.
\end{aligned}$$

最后，考虑任一 $F_0=\tilde{F}_0M_0$ $(\tilde{F}_0\in C^{k\times m})$，以及 $H\in \boldsymbol{L}_2(N)_t^k$，并记

$$F_s\triangleq F_0+(H.\ N)_s=F_0+\int_0^s H_r dN_r,\quad (0\leqslant s\leqslant t).$$

按习用的记号即 $\frac{dF}{dN}=H$. 下面要证明 $F\in\mathscr{R}_t^k$. 为此，先讨论积分方程(6.3.9)式，其中 V^F 由(6.3.11)式规定. 根据引理 6.3.2，为保证(6.3.9)式有解，还须验证 $V^{F*}\in \boldsymbol{L}_2(N)_t^k$. 事实上，由引理 5.1.6 和以上已证明的结果，得

$$\operatorname{tr}\left[\int_0^t\left(\int_0^s\left(\frac{dF}{dN}\right)_r dN_r K^*(s,\ r)\right)dN_s\left(\int_0^s K(s,\ r)dN_r\left(\frac{dF}{dN}\right)_r^*\right)\right]$$

$$=\int_0^t\left\|\int_0^s\left(\frac{dF}{dN}\right)_r dN_r K^*(s,\ r)\left(\frac{dN}{d\operatorname{tr}N}\right)_s^{\frac{1}{2}}\right\|^2 d\operatorname{tr}N_s$$

$$\leqslant\int_0^t\left\{\operatorname{tr}\left[\int_0^s\left(\frac{dF}{dN}\right)_r dN_r\left(\frac{dF}{dN}\right)_r^*\right]\right.$$

$$\times\operatorname{tr}\left[\int_0^s\left(\frac{dN}{d\operatorname{tr}N}\right)_s^{\frac{1}{2}}K(s,\ r)dN_r\right.$$

$$\left.\left.\times K^*(s,\ r)\left(\frac{dN}{d\operatorname{tr}N}\right)_s^{\frac{1}{2}}\right]\right\}d\operatorname{tr}N_s$$

$$\leqslant\operatorname{tr}\left[\int_0^t\left(\frac{dF}{dN}\right)_r dN_r\left(\frac{dF}{dN}\right)_r^*\right]$$

$$\times\operatorname{tr}\left[\int_0^t\left(\int_0^s K(s,\ r)dN_rK^*(s,\ r)\right)dN_s\right]<\infty.$$

同样可验证

$$\operatorname{tr}\left[\int_0^t R(z_0,\ x_s)(I+\Delta M_sP_s)^{-1}\right.$$

$$\left.\times dN_s(I+P_s\Delta M_s)^{-1}R(x_s,\ z_0)\right]$$

$$\leqslant\operatorname{tr}R(z_0)\operatorname{tr}\left[\int_0^t R(x_s)dM_s\right]<\infty.$$

现在，设 X^F 是(6.3.9)式在$[0,\ t]$上的一个解，即满足

$$X_s^F=(I+P_s\Delta M_s)^{-1}R(x_s,\ z_0)M_0^-F_0^*$$

$$+\int_0^s K(s,\ r)dN_r\left[\left(\frac{dF}{dN}\right)_r^*-X_r^F\right]\quad(0\leqslant s\leqslant t),$$

且 $X^{F*}\in L_2(N)_t^k$. 令

$$u_t\triangleq F_0M_0^-z_0+\int_0^t\left[\left(\frac{dF}{dN}\right)_r-X_r^{F*}\right]d\varepsilon_r,$$

则 $u_t\in\mathscr{H}(\varepsilon,\ t)^k\subset\mathscr{H}_t^k$，于是由(6.3.4)、(6.3.15)和(6.3.16)式得到：对任何 $s\in[0,\ t]$，有

$$R(u_t,\ z_s)=R\left(u_t,\ \varepsilon_s+\int_0^s dM_ry_r\right)$$

$$=F_0+\int_0^s\left[\left(\frac{dF}{dN}\right)_r-X_r^{F*}\right]dN_r$$

$$+\int_0^s R(u_r,y_r)dM_r$$

$$
\begin{aligned}
&=F_0+\int_0^s\left[\left(\frac{dF}{dN}\right)_r-X_r^{F*}\right]dN_r\\
&\quad+\int_0^s R\Big(u_r,\ R(x_r,\ z_0)M_0^- z_0\\
&\quad+\int_0^r H(r,\ \tau)\,d\varepsilon_\tau\Big)(I+\Delta M_rP_r)^{-1}dN_r\\
&=F_0+\int_0^s\left[\left(\frac{dF}{dN}\right)_r-X_r^{F*}\right]dN_r\\
&\quad+\int_0^s\Big[F_0M_0^- R(z_0,\ x_r)(I+\Delta M_rP_r)dN_r\\
&\quad+\int_0^s\left[\int_0^r\left(\left(\frac{dF}{dN}\right)_\tau-X_\tau^{F*}\right)dN_\tau K^*(r,\ \tau)\right]dN_r\\
&=F_0+\int_0^s\left(\frac{dF}{dN}\right)_r dN_r=F_s.
\end{aligned}
$$

这就证明了 $F\in\mathscr{R}_t^k$, 且 $z(F)_t=u_t\in\mathscr{H}(\varepsilon,\ t)^k$, 于是(6.3.6)式成立, 且对任意的 $F\in\mathscr{R}_t^k$, $z(F)_t\in\mathscr{H}(\varepsilon,\ t)^k$, 从而 $\mathscr{H}_t^k=\mathscr{H}(\varepsilon,\ t)^k$, $t\in[0,\ \infty)$. 这表明 $\varepsilon_{[0,\infty)}$确是 $z_{(0,\infty)}$的新息过程. u_t 的表达式又证明了(6.3.8)式, 而(6.3.7)式则易由(6.3.8)式和定理 5.3.5 的 ii)推得.

此外, 值得注意的是: 虽然我们曾在引理 6.3.2 的末尾举例说明, 积分方程(6.3.1)的解当 M 不连续时未必唯一, 但具体到积分方程(6.3.9), 即使 M 不连续(等价于 N 不连续), 其解仍然是唯一存在的. 事实上, 由(6.3.16)、(6.3.8)与(6.3.15)式,

$$
\begin{aligned}
(I+P_s dM_s)X_s^F&=R(x_s,\ z_0)M_0^-F_0\\
&\quad+\int_0^s H(s,\ r)dN_r\left[\left(\frac{dF}{dN}\right)_r^*-X_r^F\right]\\
&=R(v_s,\ z(F)_s)-R(y_s,\ z(F)_s)\\
&=S(Z^{x_t},\ F)_s.
\end{aligned}
$$

故得

$$
X_s^F=(I+P_s\Delta M_s)^{-1}S(Z^{x_t},\ F)_s(0\leqslant s\leqslant t). \quad (6.3.17)
$$

右端的唯一性即保证了 X^F 的唯一性. ▌

注记 6.3.4

i) 在定理6.3.3中, 如果假定正交增量过程 $w_{[0,\infty)}$ 的核 M 为

连续(即 $\Delta M=0$)，这等价于 $w_{[0,\infty)}$ 为二阶连续，此时新息过程 $\varepsilon_{[0,\infty)}$的核 N 就等于 M，而且定理的表达及证明还可以进一步简化. 这是一种最常见的情形. 事实上，Rozanov[7]只是给出了这种情形且 $m=1$ 时的证明，其证明的思路也完全不同于此处. 此外，他只论证了 $\varepsilon_{[0,\infty)}$是 $z_{[0,\infty)}$的新息过程，并未给出后者的再生核表示空间及其逆表示.

ii) 如在引理 5.1.4 中所指出的，若$(\boldsymbol{T}, \mathscr{T}, \mu)$为一$m$阶阵测度空间，则 $\mathrm{tr}\mu$ 为$(\boldsymbol{T}, \mathscr{T})$上的普通测度，且存在 Radon-Nikodym 导数 $0\leqslant\frac{d\mu}{d\mathrm{tr}\mu}\leqslant I_m$, a. e. $\mathrm{tr}\mu$（即除去一 $\mathrm{tr}\mu$ 零测集以外此不等式处处成立，或称对 $\mathrm{tr}\mu$ 几乎处处成立）. 于是有以下积分等式成立:

$$\mu(\boldsymbol{S})=\int_{\boldsymbol{S}}\left(\frac{d\mu}{d\mathrm{tr}\mu}\right)_t d\mathrm{tr}\mu,\ (\boldsymbol{S}\in\mathscr{T}),$$

$$\int_{\boldsymbol{T}}F_t\,d\mu G_t^*=\int_{\boldsymbol{T}}F_t\left(\frac{d\mu}{d\mathrm{tr}\mu}\right)_t G_t^*\,d\mathrm{tr}\mu.$$

对于随机积分，同样有(参见定理 5.2.4 的 viii))

$$\int_{\boldsymbol{T}}H_t\,d\mu\,\boldsymbol{z}_t=\int_{\boldsymbol{T}}H_t\left(\frac{d\mu}{d\mathrm{tr}\mu}\right)_t\boldsymbol{z}_t\,d\mathrm{tr}\mu.$$

这样就把对阵测度的积分化成了对普通测度的积分. 具体化到本章的$(\boldsymbol{T}, \mathscr{T})=([0,\infty), \mathscr{B}_{[0,\infty)})$，若 M 是$[0,\infty)$上单增的半正定 Hermite 阵函数，则对任意的 $t\in[0,\infty)$同样有

$$M_t=M_0+\int_0^t\left(\frac{dM}{d\mathrm{tr}M}\right)_s d\mathrm{tr}M_s,$$

$$\int_0^t F_s\,dM_sG_s^*=\int_0^t F_s\left(\frac{dM}{d\mathrm{tr}M}\right)_s G_s^*\,d\mathrm{tr}M_s,$$

$$\int_0^t H_s\,dM_s\boldsymbol{x}_s=\int_0^t H_s\left(\frac{dM}{d\mathrm{tr}M}\right)_s\boldsymbol{x}_s\,d\mathrm{tr}M_s.$$

其中 $\mathrm{tr}M$ 是$[0,\infty)$上的数值单增非负函数. 这种表法在本节定理的证明中已多次用过了. 此外容易看出，M 连续当且仅当 $\mathrm{tr}M$ 连续.

记 $\alpha_s=\mathrm{tr}M_s$, $D_s=\left(\frac{dM}{d\mathrm{tr}M}\right)_s$,

则有
$$M_t=M_0+\int_0^t D_s d\alpha_s.$$

利用这个方法，我们可以把重要的定理 6.3.3 翻译成更方便于应用的形式. 为简单起见，只写出 α 为连续的情形. 此外，在不致产生混淆的情况，我们把两个对 α 几乎处处相等的函数视为相等，而省去"a. e. α"的记号. 又若 J 是 $[0, t)$ 上 Borel 可测的 $C^{k\times m}$ 值函数，且有

$$\int_0^t \operatorname{tr}(J_s D_s^- J_s^*)d\alpha_s<\infty$$

成立，则记作 $J\in L_2(D^-.\alpha)_t^k$.

定理 6.3.5 设 $w_{[0,\infty)}$ 为二阶连续的 m 维正交增量过程，其核为

$$R(w_t)\triangleq M_t=M_0+\int_0^t D_s d\alpha_s,\quad (t\in[0,\ \infty)),$$

$x_{[0,\infty)}$ 为一 n 维二阶可分可测的局部正则过程，H 为一 $[0,\ \infty)$ 上 Borel 可测的 $C^{m\times n}$ 值函数，满足 $DD^-H=H$（若 D_t 恒为可逆阵，则自然成立），且对任意的 $t\in[0,\ \infty)$，有

$$\int_0^t \operatorname{tr}(H_s R(x_s) H_s^* D_s^-)d\alpha_s<\infty.$$

又设对任意的 $0\leqslant r\leqslant s<t$, $R(w_t-w_s,\ x_r)=0$. 令

$$z_t\triangleq\int_0^t H_s x_s d\alpha_s+w_t,\quad (t\in[0,\ \infty)),\qquad (6.3.18)$$

$$\varepsilon_t\triangleq z_t-\int_0^t H_s z(Z^{x_s})_s\, d\alpha_s.\qquad (6.3.19)$$

则 $\varepsilon_{[0,\infty)}$ 是以 M 为核的正交增量过程，并且是 $z_{[0,\infty)}$ 的新息过程. 此时 $z_{[0,t)}$ 的再生核表示空间及其逆再生表示分别为

$$\mathscr{R}_t^k=\left\{F=F_0+\frac{dF}{d\alpha}\cdot\alpha;\ F_0M_0^-M_0=F_0,\right.$$

$$\left.\frac{dF}{d\alpha}\ D^-D=\frac{dF}{d\alpha},\ \frac{dF}{d\alpha}\in L_2(D^-\alpha)_t^k\right\},\qquad (6.3.20)$$

$$S(F,\ G)_t=F_0M_0^-G_0^*+\int_0^t\left[\left(\frac{dF}{d\alpha}\right)_s-S(F,\ Z^{x_s})_sH_s^*\right]D_s^-\times\left[\left(\frac{dG}{d\alpha}\right)_s^*-H_sS(Z^{x_s},\ G)_s\right]d\alpha_s,$$

$$F\in\mathscr{R}_t^x,\ G\in\mathscr{R}_t^x \tag{6.3.21}$$

$$z(F)_t=F_0M_0^-z_0+\int_0^t\left[\left(\frac{dF}{d\alpha}\right)_s-S(F,\ Z^{x_s})_sH_s^*\right]D_s^-d\varepsilon_s. \tag{6.3.22}$$

其中 $S(Z^{x_s},\ F)_s$, $s\in[0,\ t]$ 满足如下的积分方程:

$$S(Z^{x_s},\ F)_s=R(x_s,\ z_0)M_0^-F_0^*+\int_0^s K(s,\ r)\left[\left(\frac{dF}{d\alpha}\right)_r^*-H_rS(Z^{x_r},\ F)_r\right]d\alpha_r, \tag{6.3.23}$$

$$K(t,\ s)\triangleq\left[\left(\frac{dZ^{x_t}}{d\alpha}\right)_s-S(Z^{x_t},\ Z^{x_s})_sH_s^*\right]D_s^-,$$

$$(0\leqslant s<t<\infty). \tag{6.3.24}$$

证　注意由条件 $DD^-H=H$, 可表

$$\int_0^t H_sx_sd\alpha_s=\int_0^t D_sD_s^-H_sx_sd\alpha_s=\int_0^t dM_s(D_s^-H_sx_s).$$

将 $D_s^-H_sx_s$ 当作定理 6.3.3 中的 x_s, 由假设条件知

$$\mathrm{tr}\left(\int_0^t R(D_s^-H_sx_s)dM_s\right)=\mathrm{tr}\left(\int_0^t D_s^-H_sR(x_s)H_s^*D_s^-D_sd\alpha_s\right)$$

$$=\int_0^t \mathrm{tr}(H_sR(x_s)H_s^*D_s^-)d\alpha_s<\infty,$$

即满足定理 6.3.3 中对 $x_{[0,\infty)}$ 所设的相应条件. 又因

$$z(Z^{D_s^-H_sx_s})=D_s^-H_sz(Z^{x_s})_s,$$

故有

$$\int_0^t H_tz(Z^{x_s})_sd\alpha_s=\int_0^t D_sD_s^-H_sz(Z^{x_s})_sd\alpha_s$$

$$=\int_0^t dM_sz(Z^{D_s^-H_sx_s})_s.$$

由此, (6.3.18)和(6.3.19)式即分别相当于(6.3.3)和(6.3.4)式, 只是将后者中的 x_s 换成了 $D_s^-H_sx_s$. 于是由定理 6.3.3, $\varepsilon_{[0,\infty)}$ 必是以 $N=M$ 为核的正交增量过程(因为 $\Delta M=0$), 而且是 $z_{[0,\infty)}$ 的

新息过程.

关于$\mathscr{R}_t^k$的界定，对照(6.3.6)式，只须注意两点：其一是条件$F_0=\widetilde{F}_0M_0$显然等价于$F_0=F_0M_0^-M_0$；其二是若$\frac{dF}{dM}\in L_2(M)_t^k$，则$\frac{dF}{d\alpha}=\frac{dF}{dM}D$，从而$\frac{dF}{d\alpha}D^-D=\frac{dF}{d\alpha}$，且$\frac{dF}{d\alpha}\in L_2(D\cdot^-\alpha)_t^k$，而反之，若后两条件成立，则$\frac{dF}{dM}=\frac{dF}{d\alpha}D^-$且$\frac{dF}{dM}\in L_2(M)_t^k$，这样就得到(6.3.20).

至于(6.3.21)和(6.3.22)式，固然可以对照(6.3.7)和(6.3.8)式求得，但在已知$\varepsilon_{[0,\infty)}$为新息过程后，更简捷的办法是直接推导。事实上，设$F\in\mathscr{R}_t^k$，则因$\mathscr{H}_t^k=\mathscr{H}(\varepsilon,t)^k$，必可表

$$z(F)_t=F_0M_0^-z_0+\int_0^t\widetilde{F}_s d\varepsilon_s,\ \ \widetilde{F}\in L_2(D\cdot\alpha)_t^k.$$

于是对任意的$s\in[0,t]$,

$$\begin{aligned}F_0+\int_0^s\left(\frac{dF}{d\alpha}\right)_r d\alpha_r=F_s&=R(z(F)_t,\ z_s)\\&=R\Big(z(F)_t,\ \varepsilon_s+\int_0^s H_r z(Z^{*r})_r d\alpha_r\Big)\\&=F_0+\int_0^s\widetilde{F}_r D_r d\alpha_r\\&\quad+\int_0^s S(F,\ Z^{*r})_r H_r^* d\alpha_r.\end{aligned}$$

对比两端即得，在$[0,t]$上

$$\widetilde{F}_s D_s=\left(\frac{dF}{d\alpha}\right)_s-S(F,\ Z^{*s})_s H_s^*\quad(\text{a}\cdot\text{e}\cdot\alpha).$$

再注意条件$\frac{dF}{d\alpha}D^-D=\frac{dF}{d\alpha}$，$DD^-H=H$　$(\text{a}\cdot\text{e}\cdot\alpha)$，即知可取

$$\widetilde{F}_s=\left[\left(\frac{dF}{d\alpha}\right)_s-S(F,\ Z^{*s})_s H_s^*\right]D_s^-,\ \ s\in[0,t].$$

由此即得(6.3.22)，从而得(6.3.21). (6.3.23)和(6.3.24)则可从(6.3.21)用Z^{*r}代F，用F代G求得. ▌

定理6.3.3和6.3.5虽然给出了形如(6.3.3)或(6.3.18)式

的 $z_{[0,\infty)}$的新息过程及其再生表示的表达式，但是它们并没有解决后者的实际计算问题．究其主要原因，则是因为对于积分方程(6.3.9)或(6.3.23)，虽然肯定了其解的存在唯一性，但却难以求出它们的解，甚至难以给出解的近似算法．为了克服这个困难，还需要对(6.3.3)或(6.3.18)式中的过程 $x_{[0,\infty)}$加以某种本质的限制，这就是下一章要讨论的连续时间线性随机系统(只限于讨论 $\alpha_t \equiv t$ 的情形)。

第 7 章

连续时间系统的线性滤波

§7.1 线性随机微分方程

定义 7.1.1 一个 n 维线性随机微分方程，通常可表为如下的积分方程形式：

$$x_t=\int_0^t A_s x_s ds+v_t, \quad (t\in[0,\infty)). \tag{7.1.1}$$

其中 $v_{[0,\infty)}$ 为二阶连续的 n 维局部正则过程．为了上式中的随机积分有意义，还要求 A 是$[0,\infty)$上 Borel 可测的 $C^{n\times n}$ 值函数，且对任意的 $t\in[0,\infty)$满足

$$\int_0^t \|A_s\|^2 ds<\infty. \tag{7.1.2}$$

在工程技术文献中，常常把(7.1.1)表成一种更为熟悉的形式，即在方程两端形式地对 t 求导，从而变成

$$\frac{dx}{dt}=A_t x_t+\frac{dv}{dt},\ x_0=v_0, \tag{7.1.3}$$

但是对于某些过程，尤其是当 $v_{[0,\infty)}$ 为最常用的正交增量过程时，$\frac{dv}{dt}$难以在普通的意义下严格定义(在广义过程的意义下则是可能的，但我们在此不加讨论)，因此我们只讨论(7.1.1)这种模式．

定理 7.1.2 用 $\mathscr{X}^n_{[0,\infty)}$ 表概率空间$(\Omega,\mathscr{F},P)$上二阶连续的 n 维局部正则过程全体．$v_{[0,\infty)}\in\mathscr{X}^n_{[0,\infty)}$，则方程(7.1.1)在 $\mathscr{X}^n_{[0,\infty)}$ 中有唯一解．

证 设 $0\leqslant t'<t''<\infty$，记 $\mathscr{X}^n_{[t',t'']}$ 为$[t',t'']$上二阶连续的 n 维正则过程全体．对任意的 $x_{[t',t'']}\in\mathscr{X}^n_{[t',t'']}$，令

$$\|x_{[t',t'']}\|_R^2 \triangleq \sup_{t'\leqslant t\leqslant t''} \operatorname{tr} R(x_t).$$

易知 $(\mathscr{X}^n_{[t',t'']}, \|\ \|_R)$ 构成一Banach 空间，在其上定义算子 A:

$$(Ax)_t \triangleq v_t + \int_{t'}^{t} A_s x_s ds, \quad (t\in[t', t'']). \tag{7.1.4}$$

则由定理 5.2.4 的 iv)，对任意 $x_{[t',t'']}$, $y_{[t',t'']}\in\mathscr{X}^n_{[t',t'']}$, 有

$$\begin{aligned}
&\|(Ax)_{[t',t'']} - (Ay)_{[t',t'']}\|_R^2\\
&= \sup_{t'\leqslant t\leqslant t''} \operatorname{tr} R\Big(\int_{t'}^{t} A_s(x_s - y_s)ds\Big)\\
&\leqslant \sqrt{n} \sup_{t'\leqslant t\leqslant t''} \Big\|R\Big(\int_{t'}^{t} A_s(x_s - y_s)ds\Big)\Big\|\\
&\leqslant \sqrt{n}\int_{t'}^{t''} \operatorname{tr}(A_s A_s^*)ds\cdot\int_{t'}^{t''} \operatorname{tr} R(x_s - y_s)ds\\
&\leqslant \sqrt{n}\ (t''-t')\Big(\int_{t'}^{t''} \|A_s\|^2 ds\Big)\|x_{[t',t'']} - y_{[t',t'']}\|_R^2.
\end{aligned} \tag{7.1.5}$$

令 $t_0=0$, 选取 $t_k\uparrow\uparrow\infty$, 使 $t_k - t_{k-1}\leqslant\frac{1}{\sqrt{n}}$, 且 $\int_{t_{k-1}}^{t_k} \|A_s\|^2 ds<1$. 令 $v_t^{(0)}\equiv v_t$. 仿(7.1.4)在 $\mathscr{X}^n_{[0,t_1]}$ 上定义算子 A_0:

$$(A_0 x)_t \triangleq v_t^{(0)} + \int_0^t A_s x_s ds, \quad (t\in[0, t_1]).$$

则由(7.1.5)知，对任意的 $x_{[0,t_1]}$, $y_{[0,t_1]}\in\mathscr{X}^n_{[0,t_1]}$ 有下式成立:

$$\|(A_0 x)_{[0,t_1]} - (A_0 y)_{[0,t_1]}\|_R \leqslant \alpha_0\|x_{[0,t_1]} - y_{[0,t_1]}\|_R,$$

其中 $0\leqslant\alpha_0<1$. 于是由熟知的压缩映象原理，在 $\mathscr{X}^n_{[0,t_1]}$ 中存在唯一的不动点使 $(A_0 x)_{[0,t_1]} = x_{[0,t_1]}$, 即方程(7.1.1)在 $\mathscr{X}^n_{[0,t_1]}$ 中有唯一解 $x_{[0,t_1]}$.

用归纳法: 设方程(7.1.1)在 $\mathscr{X}^n_{[0,t_k]}$ 中已有唯一解 $x_{[0,t_k]}$. 记

$$v_t^{(k)} \triangleq v_t + \int_0^{t_k} A_s x_s ds, \quad t\in[t_k, t_{k+1}],$$

$$(A_k x)_t \triangleq v_t^{(k)} + \int_{t_k}^{t} A_s x_s ds, \quad t\in[t_k, t_{k+1}],$$

$$x_{[t_k,t_{k+1}]}\in\mathscr{X}^n_{[t_k,t_{k+1}]}.$$

则 $(A_k x)_{t_k} = v_{t_k}^{(k)} = x_{t_k}$, 且同样由(7.1.5)可知，对任意 $x_{[t_k,t_{k+1}]}$,

$y_{[t_k,t_{k+1}]}$有下式成立:

$$\|(A_k x)_{[t_k,t_{k+1}]}-(A_k y)_{[t_k,t_{k+1}]}\|_B$$
$$\leqslant \alpha_k\|x_{[t_k,t_{k+1}]}-y_{[t_k,t_{k+1}]}\|_B.$$

其中$0\leqslant\alpha_k<1$,于是在$\mathscr{X}^n_{[t_k,t_{k+1}]}$中存在唯一的不动点使$(A_k x)_{[t_k,t_{k+1}]}=x_{[t_k,t_{k+1}]}$. 再令

$$x_{[0,t_{k+1}]}\triangleq x_{[0,t_k]}1_{[0,t_k)}+x_{[t_k,t_{k+1}]}1_{[t_k,t_{k+1}]}.$$

则显然它属于$\mathscr{X}^n_{[0,t_{k+1}]}$,且是方程(7.1.1)在$[0,\ t_{k+1}]$上的解.

为证唯一性,设$y_{[0,t_{k+1}]}$是(7.1.1)在$[0,\ t_{k+1}]$上的另一解,则由归纳法假设,必有$y_{[0,t_k]}=x_{[0,t_k]}$,且$y_{[t_k,t_{k+1}]}$必为A_k的不动点,故$y_{[t_k,t_{k+1}]}=x_{[t_k,\ t_{k+1}]}$,因而$y_{[0,t_{k+1}]}=x_{[0,t_{k+1}]}$. 由于$t_k\uparrow\infty$, 这就证明了定理. ∎

引理 7.1.3 设A满足(7.1.2)式. 用$\mathscr{C}^{n\times m}$表$[0,\ \infty)$上连续的$C^{n\times m}$值函数全体. $V\in\mathscr{C}^{n\times m}$. 则矩阵值线性微分方程

$$X_t=\int_0^t A_s X_s ds+V_t,\quad (t\in[0,\ \infty))$$

在$\mathscr{C}^{n\times m}$中有唯一解. 特别是,线性微分方程

$$\dot X_t\triangleq\frac{dX}{dt}=A_t X_t+U_t,\ t\in[0,\ \infty),\ X_{t_0}=X^0_{t_0}$$

对任何固定的t_0和$X^0_{t_0}$在$\mathscr{C}^{n\times m}$中都有唯一解. 其中U在任意有限区间$[0,\ t]$上可积.

证 证法与定理 7.1.2 类似. 只是这里在$\mathscr{C}^{n\times m}_{[t_1,t_2]}$上定义范数

$$\|X_{[t_1,t_2]}\|_B\triangleq\sup_{t_1\leqslant t\leqslant t_2}\|X_t\|.$$

由引理 5.1.6 的(5.1.1′)式容易得到与(7.1.5)式类似的不等式

$$\|(AX)_{[t_1,t_2]}-(AY)_{[t_1,t_2]}\|^2_B$$
$$\leqslant(t_2-t_1)\int_{t_1}^{t_2}\|A_s\|^2ds\|X_{[t_1,t_2]}-Y_{[t_1,t_2]}\|_B.$$

其余步骤完全一样.

至于最后一个断言,可以这样来证明:考虑两个积分形式的方程

$$X_t=\int_{t_0}^t A_sX_sds+X_{t_0}^0+\int_{t_0}^t U_sds,\ t\in[t_0,\ \infty),$$

$$X_t=-\int_t^{t_0} A_sX_sds+X_{t_0}^0-\int_t^{t_0}U_sds,\ t\in[0,\ t_0].$$

同前面的证明易知它们分别有唯一解 $X_{[t_0,\infty)}$ 和 $X_{[0,t_0]}$，而且都有 $X_{t_0}=X_{t_0}^0$. 再令

$$X_{[0,\infty)}\triangleq X_{[0,t_0]}1_{[0,t_0)}+X_{[t_0,\infty)}1_{[t_0,\infty)},$$

即得证. ∎

定义 7.1.4 设 A 满足(7.1.2)式. 考虑 n 阶阵的齐次线性微分方程

$$\frac{dX}{dt}=A_tX_t,\ (t\in[0,\ \infty)),\tag{7.1.6}$$

或简写成 $\dot{X}=AX$. 由引理 7.1.3, 对任意固定的 $s\in[0,\ \infty)$, 我们把(7.1.6)的满足 $X_s=I_n$ 的唯一解记作 $\Phi(t,s)(t\in[0,\infty))$. Φ 称为 A 的转移阵(函数), 它是$[0,\ \infty)\times[0,\ \infty)$上的 n 阶阵函数, $\Phi(s,\ s)=I$. 在本章和下一章, 我们有时不加说明地保留 Φ 作为 A 的转移阵的记号.

顺便指出, 若(7.1.6)是非时变的, 即 A_t 为 n 阶常数阵 A_0, 则 $\Phi(t,\ s)=e^{A_0(t-s)}$(矩阵函数); 若对不同的 t, A_t 之间是对乘法可交换的, 则

$$\Phi(t,\ s)=e^{\int_s^t A_rdr}.$$

定理 7.1.5 设 A 满足(7.1.2)式, Φ是 A 的转移阵. 又设 $v_{[0,\infty)}$ 是二阶连续的 n 维正交增量过程. 则

i) 对任意 t、s、$r\in[0,\ \infty)$有下式成立:

$$\Phi(t,\ s)\Phi(s,\ r)=\Phi(t,\ r),\tag{7.1.7}$$

特别有

$$\Phi(t,\ s)\Phi(s,\ t)=\Phi(t,\ t)=I.\tag{7.1.8}$$

ii) 线性随机微分方程

$$x_t=\int_0^t A_sx_sds+v_t,\ (t\in[0,\ \infty))\tag{7.1.9}$$

的唯一解可表为

$$x_t=\Phi(t,\ 0)v_0+\int_0^t\Phi(t,\ \tau)dv_\tau$$
$$=\Phi(t,\ s)x_s+\int_s^t\Phi(t,\ \tau)dv_\tau,\ (0\leqslant s\leqslant t<\infty).$$
$$(7.1.10)$$

证 i) 显然，方程(7.1.6)的满足 $X_s=X_s^0$ 的唯一解可表为 $X_t=\Phi(t,s)X_s^0(t\in[0,\infty))$. $\Phi(t,\tau)(t\in[0,\infty))$是(7.1.6)的解，它在 s 处的值为 $\Phi(s,\ \tau)$，所以 $\Phi(t,\ \tau)=\Phi(t,\ s)\Phi(s,\ \tau)$.

ii) 由定理 7.1.2，方程(7.1.9)的解是唯一的. 所以只须验证(7.1.10)是(7.1.9)的解即可. 事实上，由定理 5.5.1 和 Φ 的定义，有

$$\int_0^t A_s[\Phi(s,\ 0)v_0+\int_0^s\Phi(s,\ \tau)dv_\tau]ds$$
$$=\int_0^t\frac{d\Phi(s,\ 0)}{ds}ds\,v_0+\int_0^t\left(\int_\tau^t\frac{d\Phi(s,\ \tau)}{ds}ds\right)dv_\tau$$
$$=(\Phi(t,\ 0)-I)v_0+\int_0^t(\Phi(t,\ \tau)-I)dv_\tau$$
$$=\Phi(t,\ 0)v_0+\int_0^t\Phi(t,\ \tau)dv_\tau-v_t.$$

故(7.1.10)的中间一式确是(7.1.9)的解. 又由(7.1.7)，有

$$\Phi(t,\ 0)v_0+\int_0^t\Phi(t,\ \tau)dv_\tau$$
$$=\Phi(t,\ s)\left[\Phi(s,\ 0)v_0+\int_0^s\Phi(s,\tau)dv_\tau\right]+\int_s^t\Phi(t,\ \tau)dv_\tau$$
$$=\Phi(t,\ s)x_s+\int_s^t\Phi(t,\ \tau)dv_\tau.\quad\blacksquare$$

§7.2 线性系统输出的再生表示及滤波

定义 7.2.1 如下的联立方程确定一个连续时间的$(n,\ m)$阶线性随机系统：

$$\begin{cases} x_t = \int_0^t A_s x_s ds + v_t; \\ z_t = \int_0^t H_s x_s ds + w_t, \end{cases} \quad t \in [0, \infty). \tag{7.2.1}$$

其中 $u_{[0,\infty)} = \left(u_t \triangleq \begin{pmatrix} v_t \\ w_t \end{pmatrix}, \ t \in [0, \infty) \right)$ 为 $n+m$ 维(局部正则的)正交增量过程，并设其核可表为

$$R(u_t) = R(u_0) + \int_0^t \begin{pmatrix} C_s & B_s \\ B_s^* & D_s \end{pmatrix} ds. \tag{7.2.2}$$

又设 A 和 H 分别是 $[0, \infty)$ 上 Borel 可测的 $C^{n\times n}$ 和 $C^{m\times n}$ 值函数(不熟悉测度论的读者可以认为它们是连续函数)，满足条件 $DD^- H = H$(a.e.)(当 D_s 恒为可逆阵时，条件自然成立)，且对任意的 $t \in [0, \infty)$ 有下式成立：

$$\int_0^t \|A_s\|^2 ds < \infty, \quad \int_0^t \mathrm{tr}(H_s^* D_s^- H_s) ds < \infty. \tag{7.2.3}$$

系统(7.2.1)与离散时间的线性系统(4.1.1)是对应的，它的第一个方程称为*动态方程*，$x_{[0,\infty)}$ 称为*动态过程*，x_t 称为*系统在 t 时的状态*，$z_{[0,\infty)}$ 称为*系统的量测或输出过程*，$v_{[0,\infty)}$ 和 $w_{[0,\infty)}$ 分别称为*动态噪声*和*量测噪声*。换句话说，系统的状态是不能直接量测到的，我们只能通过含有噪声的"部分"观测值 $z_{[0,\infty)}$ 来对 $x_{[0,\infty)}$ 作出估计，这就是滤波问题.

注意，从 $u_{[0,\infty)}$ 为正交增量过程这一假定及(7.2.2)式立即推知：$v_{[0,\infty)}$ 和 $w_{[0,\infty)}$ 分别是二阶连续的 n 维和 m 维正交增量过程，它们的核分别为

$$R(v_t) = R(v_0) + \int_0^t C_s ds, \ R(w_t) = R(w_0) + \int_0^t D_s ds,$$
$$(t \in [0, \infty)),$$

并且对任意的 $0 \leqslant s \leqslant t < \infty$ 有下式成立：

$$R(v_t - v_s, \ w_s) = 0, \ R(w_t - w_s, \ v_s) = 0,$$
$$R(v_t, \ w_t) = R(v_0, \ w_0) + \int_0^t B_s ds.$$

又由(7.2.1)显然有 $x_0 = v_0$, $z_0 = w_0$.

定理 7.2.2 设 $z_{[0,\infty)}$ 是系统(7.2.1)的输出过程. 对任意的 $t\in[0,\infty)$. 令

$$y_t\triangleq z(Z^{x_t})_t, \tag{7.2.4}$$

$$P_t\triangleq R(x_t-y_t)=R(x_t)-R(y_t)=R(x_t)-S(Z^{x_t})_t, \tag{7.2.5}$$

$$\varepsilon_t\triangleq z_t-\int_0^t A_s y_s ds. \tag{7.2.6}$$

则 $\varepsilon_{[0,\infty)}$ 是二阶连续的 m 维正交增量过程, 并且是 $z_{[0,\infty)}$ 的新息过程, 其核与 $w_{[0,\infty)}$ 的核相同, $\varepsilon_0=z_0=w_0$.

此时, $y_{[0,\infty)}$ 满足如下的线性随机微分方程:

$$y_t=R(x_0,\ z_0)R(z_0)^-z_0+\int_0^t A_s y_s ds+\int_0^t (P_sH_s^*+B_s)D_s^-d\varepsilon_s,$$

$$(t\in[0,\infty)). \tag{7.2.7}$$

P 满足如下的 n 阶阵的 Ricatti 型(二次)微分方程:

$$\dot{P}=AP+PA^*+C-(PH^*+B)D^-(HP+B^*), \tag{7.2.8}$$

$$P_0=R(x_0)+R(x_0,\ z_0)R(z_0)^-R(z_0,\ x_0). \tag{7.2.9}$$

其中 $\dot{P}_t\triangleq \dfrac{dP}{dt}$, 而 C、D、B 由(7.2.2)式给出.

若用 $\mathscr{C}_{at}$ 记 $[0,t]$ 上的绝对连续函数(即可表为其导函数的不定积分的函数)全体, 则此时 $z_{[0,t]}$ 的再生核表示空间及其逆表示分别为

$$\mathscr{R}_t^k=\{F\in\mathscr{C}_{at}^{k\times m}:\ F_0R(z_0)^-R(z_0)=F_0,\ \dot{F}D^-D=\dot{F},$$

$$\int_0^\infty \operatorname{tr}(\dot{F}_sD_s^-\dot{F}_s^*)ds<\infty\}. \tag{7.2.10}$$

$$S(F,\ G)_t=F_0R(z_0)^-G_0^*+\int_0^t(\dot{F}_s-Y_s^{F*}H_s^*)D_s^-(\dot{G}_s^*$$

$$-H_sY_s^G)ds,\ (F\in\mathscr{R}_t^k,\ G\in\mathscr{R}_t^l). \tag{7.2.11}$$

$$z(F)_t=F_0R(z_0)^-z_0+\int_0^t(\dot{F}_s-Y_s^{F*}H_s^*)D_s^-d\varepsilon_s,\ (F\in\mathscr{R}_t^k) \tag{7.2.12}$$

其中 $Y_s^F\triangleq S(Z^{x_s},\ F)_s$ 在 $[0,t]$ 上满足如下 $n\times k$ 阵的线性微分方程:

$$\dot{Y}^F = AY^F + (PH^* + B)D^-(\dot{F}^* - HY^F), \qquad (7.2.13)$$

$$Y_0^F = R(x_0, z_0)R(z_0)^- F_0^*. \qquad (7.2.14)$$

证 本定理关于 $\varepsilon_{[0,\infty)}$ 是新息过程的结论以及(7.2.10)～(7.2.12)式不过是定理 6.3.5 当 $\alpha_t \equiv t$ 时这种特殊情形的改写(注意那里的 $M_0 = R(w_0) = R(z_0)$). 所不同的只是：这时 $x_{[0,\infty)}$ 是(7.2.1)中的随机微分方程的解，因此是二阶连续的，从而由定理 6.1.4，是二阶可分可测的. 此时由 $R(x_t)$ 在 $[0, \infty)$ 上的连续性及(7.2.3)式还可推出

$$\int_0^t \mathrm{tr}(H_s R(x_s) H_s^* D_s^-) ds = \int_0^t \mathrm{tr}(R(x_s) H_s^* D_s^- H_s) ds$$

$$\leqslant \sup_{0 \leqslant s \leqslant t} \|R(x_s)\| \int_0^t \mathrm{tr}(H_s^* D_s^- H_s) ds < \infty.$$

此外，从 $u_{[0,\infty)}$ 为正交增量过程又易知，对任意的 $0 \leqslant r \leqslant s < t$，$R(w_t - w_s, x_r) = 0$，所以定理 6.3.5 的假定条件全部满足.

剩下要证明的是(7.2.7)、(7.2.8)和(7.2.13)式.

由系统(7.2.1)的假设条件与定理 7.1.5，容易对任意 $0 \leqslant s \leqslant t < \infty$ 依次算出下面一系列等式：

$$x_t = \Phi(t, s)x_s + \int_s^t \Phi(t, r) dv_r,$$

$$\begin{aligned}
R(x_t, x_s) &= \Phi(t, s)R(x_s) \\
&= \Phi(t, s)R\left(\Phi(s, 0)x_0 + \int_0^s \Phi(s, r) dv_r\right) \\
&= \Phi(t, 0)R(x_0)\Phi^*(s, 0) + \int_0^s \Phi(t, r) C_r \Phi^*(s, r) dr.
\end{aligned} \qquad (7.2.15)$$

$$\begin{aligned}
Z_t^{x_s} \triangleq R(x_t, z_s) &= \Phi(t, s)R(x_s, z_s) \\
&= \Phi(t, s)R\left(x_s, \int_0^s H_r x_r dr + w_s\right) \\
&= \Phi(t, s)\left(\int_0^s \Phi(s, r)R(x_r)H_r^* dr \right. \\
&\quad \left. + \Phi(s, 0)R(x_0, z_0) + \int_0^s \Phi(s, r)B_r dr\right)
\end{aligned}$$

$$=\Phi(t,0)R(x_0,z_0)+\int_0^s\Phi(t,r)(R(x_r)H_r^*+B_r)dr. \tag{7.2.16}$$

$$\frac{dZ_s^{x_t}}{ds}=\Phi(t,s)(R(x_s)H_s^*+B_s), \tag{7.2.17}$$

$$\begin{aligned}S(Z^{x_t},Z^{x_s})_s&=R(z(Z^{x_t})_s,z(Z^{x_s})_s)\\&=R(x_t,y_s)=\Phi(t,s)R(x_s,y_s)\\&=\Phi(t,s)R(y_s).\end{aligned} \tag{7.2.18}$$

将(7.2.17)和(7.2.18)代入定理 6.3.5 中的(6.3.24)式即得

$$\begin{aligned}K(t,s)&\triangleq\left[\frac{dZ_s^{x_t}}{ds}-S(Z^{x_t},Z^{x_s})_sH_s^*\right]D_s^-\\&=\Phi(t,s)[(B(x_s)-R(y_s))H_s^*+B_s]D_s^-\\&=\Phi(t,s)(P_sH_s^*+B_s)D_s^-.\end{aligned}$$

再代入(6.3.23)式得到，对 $F\in\mathscr{R}_t^*$, $s\in[0,t]$，有

$$\begin{aligned}Y_s^F\triangleq S(Z^{x_s},F)_s&=\Phi(s,0)R(x_0,z_0)R(z_0)^-F_0^*\\&+\int_0^s\Phi(s,r)(P_rH_r^*+B_r)D_r^-(\dot F_r^*-H_rY_r^F)dr.\end{aligned} \tag{7.2.19}$$

两边对 s 求导，由 Φ 的定义立刻推出

$$\dot Y_s^F=A_sY_s^F+(P_sH_s^*+B_s)D_s^-(\dot F_s^*-H_sY_s^F),$$

且其初值为 $Y_0^F=R(x_0,z_0)B(z_0)^-F_0^*$，这就是(7.2.13)和(7.2.14)。

现在，在(7.2.19)式中用 Z^{x_t} 代替 F 并注意(7.2.17)和(7.2.18)得到

$$\begin{aligned}R(y_s)\Phi^*(t,s)&=S(Z^{x_s},Z^{x_t})_s\\&=\Phi(s,0)R(x_0,z_0)R(z_0)^-R(z_0,x_0)\Phi^*(t,0)\\&+\int_0^s\Phi(s,r)(P_rH_r^*+B_r)D_r^-(H_rP_r+R_r^*)\Phi^*(t,r)dr.\end{aligned} \tag{7.2.20}$$

令 $t=s$ 并对 s 求导即有

$$\frac{dR(y_s)}{ds}=A_sR(y_s)+R(y_s)A_s^*$$

$$+(P_sH_s^*+B_s)D_s^-(H_sP_s+B_s^*).$$

又在(7.2.15)式中令 $t=s$ 并对 s 求导，得

$$\frac{dR(x_s)}{ds}=A_sR(x_s)+R(x_s)A_s^*+C_s.$$

从后式减前式，即推出 P 满足的 Ricatti 型方程(7.2.8)，且显然有形如(7.2.9)的初值.

最后，因为 $y_t\triangleq z(Z^{x_t})_t$，在(7.2.12)式中用 Z^{x_t} 代 F，并注意(7.2.17)和(7.2.18)式即得

$$y_t=\Phi(t,\ 0)R(x_0,\ z_0)R(z_0)^-z_0$$
$$+\int_0^t\Phi(t,\ s)(P_sH_s^*+B_s)D_s^-d\varepsilon_s.$$

根据定理 7.1.5 的 ii). 此即表明 $y_{[0,\infty)}$ 满足线性随机微分方程(7.2.7). ▌

注记 7.2.3 定理 7.2.2 与定理 6.3.5 的一个显著区别是把积分方程变成了微分方程. 虽然像(7.2.8)和(7.2.13)这样的二次和线性常微分方程一般还是无法求得其解析形式的解的(除非定常情形)，但当系数矩阵为连续函数时，却可以用数值方法求其近似解. 例如，如果已经由(7.2.8)求得了 P_t 的近似解，令 $K_t\triangleq(P_tH_t^*+B_t)D_t^-$，则可对新息 $\varepsilon_{[0,\infty)}$ 作如下的近似递推计算(当 $t+\delta$ 时的量测 $z_{t+\delta}$ 到来时)：由(7.2.6)，有

$$\varepsilon_{t+\delta}=\varepsilon_t+z_{t+\delta}-z_t-\int_t^{t+\delta}H_sy_sds$$
$$\sim\varepsilon_t+z_{t+\delta}-z_t-\delta H_ty_t.$$

再由(7.2.7)，有

$$y_{t+\delta}=y_t+\int_t^{t+\delta}A_sy_sds+\int_t^{t+\delta}K_sd\varepsilon_s$$
$$\sim(I+\delta A_t)y_t+K_t(\varepsilon_{t+\delta}-\varepsilon_t).$$

仿此还可由(7.2.13)、(7.2.11)和(7.2.12)求得任意 $S(F,\ G)_t$ 及 $z(F)_t$ 的近似递推计算.

下面利用定理 7.2.2 的结果来解决系统(7.2.1)的滤波问题.

定义 7.2.4 类似于记号 4.0.1 中的规定，对 τ、$t\in[0,\infty)$，我们用 $\hat{x}_{\tau|t}$ 表 x_τ 基于量测 $z_{[0,t]}$ 的线性最小方差预报(见第二章 § 2.1)，即

$$\hat{x}_{\tau|t}\triangleq\pi(x_\tau|z_{[0,t]}),$$

并记 $$P_{\tau|t}\triangleq R(x_\tau-\hat{x}_{\tau|t})=R(x_\tau)-R(\hat{x}_{\tau|t}),$$

又简记 $\hat{x}_t\triangleq\hat{x}_{t|t}$，$P_t\triangleq P_{t|t}$. 对 $t=\tau$、$t<\tau$、$t>\tau$ 三种情形，$\hat{x}_{\tau|t}$ 分别称为 x_τ 的滤波、$\tau-t$ 步预测(或外推)，$t-\tau$ 步平滑(或内插)，而 $P_{\tau|t}$ 则分别称为它们的误差方差阵.

定理 7.2.5 设线性随机系统由(7.2.1)式确定，并满足定义 7.2.1 中假定的条件. 则 $x_{[0,\infty)}$ 的滤波、预测与平滑及其误差方差阵分别满足如下的微分方程或积分表示式：

$$\varepsilon_t=z_t-\int_0^t H_s\hat{x}_s ds-Ew_t, \tag{7.2.21}$$

$$\begin{aligned}\hat{x}_t=&R(x_0,\ z_0)R(z_0)^-(z_0-Ez_0)+\int_0^t A_s\hat{x}_s ds\\&+\int_0^t(P_sH_s^*+B_s)D_s^- d\varepsilon_s+Ev_t,\end{aligned} \tag{7.2.22}$$

$$\begin{aligned}&\frac{dP_t}{dt}=A_tP_t+P_tA_t^*-(P_tH_t^*+B_t)D_t^-(H_tP_t+B_t^*)+C_t,\\&P_0=R(x_0)-B(x_0,\ z_0)R(z_0)^-R(z_0,\ x_0).\end{aligned} \tag{7.2.23}$$

当 $\tau>t$ 时，有

$$\hat{x}_{\tau|t}=\hat{x}_t+\int_t^\tau A_s\hat{x}_{s|t}ds+Ev_\tau-Ev_t, \tag{7.2.24}$$

$$\frac{dP_{\tau|t}}{d\tau}=A_\tau P_{\tau|t}+P_{\tau|t}A_\tau^*+C_\tau,\ P_{t|t}=P_t. \tag{7.2.25}$$

当 $\tau<t$ 时，设 $Q_{s\tau}$ 为如下的齐次线性微分方程的解：

$$\frac{dQ_{s\tau}}{ds}=[A_s-(P_sH_s^*+B_s)D_s^-H_s]Q_{s\tau},\ (s>\tau), \tag{7.2.26}$$

$$Q_{\tau\tau}=P_\tau.$$

则有

$$\hat{x}_{\tau|t}=\hat{x}_\tau+\int_\tau^t Q_{s\tau}^*H_s^*D_s^-d\varepsilon_s, \tag{7.2.27}$$

$$P_{\tau|t}=P_\tau-\int_\tau^t Q_{s\tau}^* H_s^* D_s^- H_s Q_{s\tau}\,ds. \tag{7.2.28}$$

证　首先注意，由定理2.1.2，$\hat{x}_{\tau|t}$与$z(Z^{x_\tau})_t$只是相差一个由均值决定的常值向量，从而滤波误差方差阵P_t与(7.2.5)中的P_t是相同的．其次，在$v_{[0,\infty)}$与$w_{[0,\infty)}$的均值函数都是已知的假定下，在证明过程中，不妨设Ev_t，Ew_t都恒等于0，从而Ex_t、Ez_t也恒为0．因为否则，可考虑与(7.2.1)等价的如下系统：

$$x_t-Ex_t=\int_0^t A_s(x_s-Ex_s)ds+v_t-Ev_t,$$

$$z_t-Ez_t=\int_0^t H_s(x_s-Ex_s)ds+w_t-Ew_t.$$

这样，(7.2.21)、(7.2.22)、(7.2.23)式就分别相当于定理7.2.2中的(7.2.6)、(7.2.7)、(7.2.8)与(7.2.9)式．唯一要注意的是：在均值不为0的情形，(7.2.21)给出的ε_t实际是z_t-Ez_t的新息，但我们仍采用与z_t的新息相同的记号．显然这并不影响问题的实质．

现在考虑$\tau>t$的情况：由(7.1.10)式，对任意$s\in[0,t]$，有

$$\begin{aligned}R(\hat{x}_{\tau|t},z_s)&=R(x_\tau,z_s)\\&=R\left(\Phi(\tau,t)x_t+\int_t^\tau\Phi(\tau,r)dv_r,\ z_s\right)\\&=\Phi(\tau,t)R(x_t,z_s)=R(\Phi(\tau,t)\hat{x}_t,z_s).\end{aligned}$$

因此（在均值为0的假定下）有

$$\hat{x}_{\tau|t}=\Phi(\tau,t)\hat{x}_t\ (\tau\geqslant t).$$

再根据(7.1.10)和(7.1.9)式，改以t为起点，并用$\hat{x}_t 1_{[t,\infty)}$代替那里的$v_{[0,\infty)}$，即得

$$\hat{x}_{\tau|t}=\hat{x}_t+\int_t^\tau A_s\hat{x}_{s|t}ds.$$

这相当于(7.2.24)式．又由(7.2.27)式得

$$R(\hat{x}_{\tau|t})=\Phi(\tau,t)R(\hat{x}_t)\Phi^*(\tau,t).$$

由(7.1.10)式得

$$R(x_\tau)=\Phi(\tau,t)R(x_t)\Phi^*(\tau,t)+\int_t^\tau\Phi(\tau,r)C_r\Phi^*(\tau,r)dr.$$

由后式减前式再对 τ 求导，即证得(7.2.25)式.

最后考虑 $\tau<t$ 的情况：由(7.2.12)式，用 Z^{x_τ} 代替其中的 F，并暂记

$$Y_s^\tau \triangleq S(Z^{x_s}, Z^{x_\tau})_s, \quad (s\in[0, \infty)),$$

得到

$$\begin{aligned}\hat{x}_{\tau|t} &= z(Z^{x_\tau})_t \\ &= \Phi(\tau, 0)R(x_0, z_0)R(z_0)^- z_0 + \int_0^t \left(\frac{dZ_s^{x_\tau}}{ds} - Y_s^{\tau *}H_s^*\right)D_s^- d\varepsilon_s \\ &= \hat{x}_\tau + \int_\tau^t \left(\frac{dZ_s^{x_\tau}}{ds} - Y_s^{\tau *}H_s^*\right)D_s^- d\varepsilon_s. \end{aligned} \tag{7.2.29}$$

另一方面，当 $s\geqslant\tau$ 时，由(7.2.1)的假定条件及(7.2.15)式容易算出

$$\begin{aligned}Z_s^{x_\tau} &\triangleq R(x_\tau, z_s) \\ &= R\left(x_\tau, z_\tau + \int_\tau^s H_r x_r dr + w_s - w_\tau\right) \\ &= R(x_\tau, z_\tau) + R(x_\tau)\int_\tau^s \Phi^*(r, \tau)H_r^* dr.\end{aligned}$$

在两端对 s 求导，得

$$\frac{dZ_s^{x_\tau}}{ds} = R(x_\tau)\Phi^*(s, \tau)H_s^*, (s>\tau). \tag{7.2.30}$$

又由(7.2.13)，Y^τ 在 $[0, \infty)$ 上满足微分方程

$$\frac{dY_s^\tau}{ds} = A_s Y_s^\tau + (P_s H_s^* + B_s)D_s^-\left[\left(\frac{dZ_s^{x_\tau}}{ds}\right)^* - H_s Y_s^\tau\right],$$

再将(7.2.30)代入，即知 Y^τ 在 $[\tau, \infty)$ 上满足方程

$$\frac{dY_s^\tau}{ds} = A_s Y_s^\tau + (P_s H_s^* + B_s)D_s^- H_s(\Phi(s, \tau)R(x_\tau) - Y_s^\tau). \tag{7.2.31}$$

现在，令 $\qquad Q_{s\tau} \triangleq \Phi(s, \tau)R(x_\tau) - Y_s^\tau, \ (s\geqslant\tau).$

注意 $\qquad \dfrac{d\Phi(s, \tau)R(x_\tau)}{ds} = A_s\Phi(s, \tau)R(x_\tau).$

由此式减去(7.2.31)，即知 $(Q_{s\tau}, s>\tau)$ 必满足方程(7.2.26)．显然 $Q_{\tau\tau} = P_\tau$，且由(7.2.31)知

$$\frac{dZ_s^{x_s}}{ds}-Y_s^{\tau*}H_s^*=Q_{s\tau}^*H_s^*. \tag{7.2.32}$$

将此式代入(7.2.29)式，即证得(7.2.27)式，并由此推出(7.2.28)式。 ▮

§7.3 含未知输入及非积累量测的情形

同第四章§4.2对应，现在考虑连续时间线性系统(7.2.1)含有未知输入时的滤波问题. 在定理7.2.5中考虑(7.2.1)的滤波问题时，实际上假定了系统不包含任何未知参数，即系数矩阵 A 和 H，噪声的核矩阵 C、D、B，初假方差与协方差阵 $R(x_0)$、$R(z_0)$、$R(x_0, z_0)$ 以及均值函数 $Ev_{[0,\infty)}$、$Ew_{[0,\infty)}$ 都是已知的. 这样，定理7.2.5列举的各种公式才能得以实现. 现在进一步讨论上述均值函数含有未知的线性回归系数向量 $\theta\in C^k$ 的情形或称之为"含未知输入"的情形(因为 $u_{[0,\infty)}$、$v_{[0,\infty)}$ 的均值函数可看作是系统(7.2.1)的常值输入).

假设与定义 7.3.1 设系统(7.2.1)中的 $n+m$ 维正交增量过程 $u_{[0,\infty)}$ 的均值函数含有 k 维未知线性回归参数 θ:

$$u_{[0,\infty)}\triangleq\begin{pmatrix}v_{[0,\infty)}\\w_{[0,\infty)}\end{pmatrix},\ E_\theta u_t=\begin{pmatrix}F_t^*\\G_t^*\end{pmatrix}\theta\quad(t\in[0,\infty)),$$

且设对任意 $t\in[0,\infty)$，有 $(F, G)\in\mathscr{R}(u, t)^k$. 因此，由定理1.2.11知 $u_{[0,\infty)}$ 为局部正则过程. 又因为显然有 $\mathscr{H}((x, z), t)\subset\mathscr{H}(u, t)$，所以由定义1.2.2又知 $x_{[0,\infty)}$ 和 $z_{[0,\infty)}$ 都是局部正则过程.

为了与定义7.2.4中的记号有所区别，同第4章§4.2，用 $\check{x}_{\tau|t}$ 记 x_τ 基于量测 $z_{[0,t]}$ 的 Gauss-Markov 估计(简称GM滤波，参见定义2.3.1)，且记 $\Pi_{\tau|t}\triangleq R(x_\tau-\check{x}_{\tau|t})$. 又若 θ 有已知的先验一、二阶矩: $E_{a\theta}\theta=\mu_{a\theta}$, $R_{a\theta}(\theta)=\Gamma_{a\theta}$，则用 $\check{\check{x}}_{\tau|t}$ 记其相应的线性 Bayes 估计(简称LB滤波，参见定义2.3.4)，且记

$$Q_{\tau|t}\triangleq E_{a\theta}E_\theta(x_\tau-\check{\check{x}}_{\tau|t})(x_\tau-\check{\check{x}}_{\tau|t})^*.$$

根据定理 7.2.2(或定理 6.2.4)，在上述假定下，F 和 G 分别是$[0, \infty)$上的 $C^{k\times n}$ 和 $C^{k\times m}$ 值绝对连续函数，即

$$F_t=F_0+\int_0^t \dot{F}_s ds,\ G_t=G_0+\int_0^t \dot{G}_s ds,\ (t\in[0, \infty)),$$

且满足
$$F_0R(x_0)^-R(x_0)=F_0,\ \dot{F}C^-C=\dot{F},$$
$$G_0R(z_0)^-R(z_0)=G_0,\ \dot{G}D^-D=\dot{G}.$$

注意由(7.2.1)式，对任意 $t\in[0, \infty)$与 $\theta\in C^k$ 有

$$E_\theta x_t=\int_0^t A_s(E_\theta x_s)ds+F_0^*\theta+\left(\int_0^t \dot{F}_s^*ds\right)\theta,$$

$$E_\theta z_t=\int_0^t H_s(E_\theta x_s)ds+G_0^*\theta+\left(\int_0^t \dot{G}_s^*ds\right)\theta.$$

故若 J^* 和 K^* 是下列线性微分方程的解：

$$\dot{J}^*=AJ^*+\dot{F}^*,\ J_0^*=F_0^*, \tag{7.3.1}$$

$$\dot{K}^*=HJ^*+\dot{G}^*,\ K_0^*=G_0^*. \tag{7.3.2}$$

则显然有

$$E_\theta x_t=J_t^*\theta;\ E_\theta z_t=K_t^*\theta,\ t\in[0, \infty),\ \theta\in C^k. \tag{7.3.3}$$

此外，在所讨论的问题中，很有用的一种特殊情形是：

$$F_0\neq 0(\text{或 } k=n \text{ 而 } F_0=I_n)\text{但 } \dot{F}=0 \text{ 且 } G=0.$$

这种情形称为*缺初值情形*(即系统的初始状态均值为未知)。

定理 7.3.2 设系统(7.2.1)满足假定 7.3.1 的条件，K由(7.3.1)和(7.3.2)定出，则对任意 $t\in[0, \infty)$，$K\in\mathscr{R}_t^*$，且此时 $z(K)_t$ 与 $S(K)_t$ 有如下的积分表达：

$$z(K)_t=G_0R(z_0)^-z_0+\int_0^t (L_s^*H_s^*+\dot{G}_s)D_s^-dz_s, \tag{7.3.4}$$

$$S(K)_t=G_0R(z_0)^-G_0^*+\int_0^t (L_s^*H_s^*+\dot{G}_s)D_s^-(H_sL_s+\dot{G}_s^*)ds. \tag{7.3.5}$$

其中 L 为如下微分方程的解：

$$\dot{L}=[A-(PH^*+B)D^-H]L+\dot{F}^*-(PH^*+B)D^-\dot{G}^*,$$
$$L_0=F_0^*-B(x_0,\ z_0)B(z_0)^-G_0^*. \tag{7.3.6}$$

而 P 由方程(7.2.23)解出。

证 首先，由于 $z_{[0,\infty)}$ 是局部正则过程，且 $E_\theta z_t = K_t^*\theta$，$\theta\in C^k$，所以由定理 1.2.11 知 $K\in\mathscr{R}^*$.

其次，由定理 7.2.2 的(7.2.12)和(7.2.13)式，有

$$z(K)_t = K_0R(z_0)^- z_0 + \int_0^t (\dot{K}_s - Y_s^{k*}H_s^*)D_s^- dz_s, \tag{7.3.7}$$

其中 $Y_s^k \triangleq S(Z^{x_s}, K)_s$ 满足微分方程($\dot{K}^*$ 用(7.3.2)式代入)：

$$\begin{aligned}\dot{Y}^k &= AY^k + (PH^* + B)D^-(\dot{K}^* - HY^k)\\ &= AY^k + (PH^* + B)D^-[H(J^* - Y^k) + \dot{G}^*],\end{aligned}$$

且 $\quad Y_0^k = R(x_0, z_0)R(z_0)^- K_0^* = R(x_0, z_0)R(z_0)^- G_0^*.$

现在，令 $L\triangleq J^* - Y^k$，则有 $\dot{K}^* - HY^k = HL + \dot{G}^*$. 代入(7.3.7)即得(7.3.4)，从而得(7.3.5). 又从(7.3.1)式减去上式，即得方程(7.3.6)，且

$$L_0 = J_0^* - Y_0^k = F_0^* - R(x_0, z_0)R(z_0)^- G_0^*. \quad \blacksquare$$

定理 7.3.3 设同定理 7.3.2. 若存在 $t_0\in[0,\infty)$，使当 K 局限于 $[0, t_0]$ 上时，其诸行是线性无关的向量函数(即 $\theta^*K_t = 0$，$t\in[0, t_0]$ 蕴含 $\theta = 0$)，则当 $t\geqslant t_0$，$\tau\in[0,\infty)$ 时，x_τ 的 GM 估计 $\check{x}_{\tau|t}$ 恒存在. 记 $L_{\tau t}\triangleq J_\tau^* - S(Z^{x_\tau}, K)_t$，则 $\check{x}_{\tau|t}$ 及其误差方差阵 $\Pi_{\tau|t}$ 分别由下式算出：

$$\check{x}_{\tau|t} = \hat{x}_{\tau|t} + L_{\tau t}S(K)_t^{-1}z(K)_t, \tag{7.3.8}$$

$$\Pi_{\tau|t} = P_{\tau|t} + L_{\tau t}S(K)_t^{-1}L_{\tau t}^*. \tag{7.3.9}$$

而 $L_{\tau t}$ 分别满足如下的微分方程与积分表达式：

$$L_{tt} = L_t,\quad \frac{dL_{\tau t}}{d\tau} = A_\tau L_{\tau t} + \dot{F}_\tau^*,\quad \tau > t, \tag{7.3.10}$$

$$L_{\tau t} = L_\tau - \int_\tau^t Q_{s\tau}^* H_s^* D_s^- (H_s L_s + \dot{G}_s^*)ds,\quad \tau < t. \tag{7.3.11}$$

在以上诸式中，$\hat{x}_{\tau|t}$、$P_{\tau|t}$、$Q_{s\tau}$ 由定理 7.2.5 算出(但那里的 Ev_t，Ew_t 都用 0 代替)，$z(K)_t$、$S(K)_t$、L_t 由定理 7.3.2 算出.

证 首先由定理 2.2.2，K 局限于 $[0, t_0]$ 上的行函数线性无关与 $S(K)_{t_0}$ 为可逆阵是互为充分必要条件的，从而当 $t\geqslant t_0$ 时，$S(K)_t$ 恒为可逆阵. 此时，由系 2.3.3，x_τ 基于 $z_{[0,t]}$ 的 Gauss-

Markov 估计 $\breve{x}_{\tau|t}$ 一定存在，且

$$\breve{x}_{\tau|t}=z(Z^{x_\tau})_t+(J_\tau^*-S(Z^{x_\tau},\ K)_t)z(K)_t.$$

而这就是(7.3.8)式，同理得(7.3.9)式.

为证(7.3.10)与(7.3.11)式，先考虑 $\tau\geqslant t$ 的情形：此时

$$\begin{aligned}S(Z^{x_\tau},\ K)_t&=R(z(Z^{x_\tau})_t,\ z(K)_t)=R(x_\tau,\ z(K)_t)\\&=R(\Phi(\tau,\ t)x_t+\int_t^\tau\Phi(\tau,\ r)dv_r,\ z(K)_t)\\&=\Phi(\tau,\ t)R(x_t,\ z(K)_t)=\Phi(\tau,\ t)S(Z^{x_t},K)_t\\&\triangleq\Phi(\tau,\ t)Y_t^k.\end{aligned}$$

因此有
$$L_{\tau t}=J_\tau^*-\Phi(\tau,\ t)Y_t^k,$$

从而由定理 7.3.2 证明中的定义，有

$$L_{tt}=J_t^*-Y_t^k\triangleq L_t.$$

又由(7.3.1)得

$$\begin{aligned}\frac{dL_{\tau t}}{d\tau}&=A_\tau J_\tau^*+\dot{F}_\tau^*-A_\tau\Phi(\tau,\ t)Y_t^k\\&=A_\tau L_{\tau t}+\dot{F}_\tau^*,\quad(\tau>t).\end{aligned}$$

于是证得(7.3.10).

再考虑 $\tau<t$ 的情形：由(7.2.11)式，并注意(7.2.32)式及定理 7.3.2 的证明，即得

$$\begin{aligned}&\frac{dS(Z^{x_\tau},\ K)_t}{dt}\\&=\left(\frac{dZ_t^{x_\tau}}{dt}-S(Z^{x_\tau},\ Z^{x_t})_tH_t^*\right)D_t^-(\dot{K}_t^*-H_tS(Z^{x_t},\ K)_t)\\&\triangleq\left(\frac{dZ_t^{x_\tau}}{dt}-Y^{\tau *}H_t^*\right)D_t^-(\dot{K}_t^*-H_tY_t^k)\\&=Q_{t\tau}^*H_t^*D_t^-(H_tL_t+\dot{G}_t^*)\quad(t>\tau).\end{aligned}$$

于是有
$$\frac{dL_{\tau t}}{dt}=-Q_{t\tau}^*H_t^*D_t^-(H_tL_t+\dot{G}_t^*)\quad(t>\tau).$$

将上式两端在$[\tau,\ t]$上积分即证得(7.3.11)式.

定理 7.3.4 设同定理 7.3.2. 又设 θ 有已知的先验一、二阶矩：$E_{a\theta}=\mu_{a\theta}^x$，$R_{a\theta}(\theta)=\Gamma_{a\theta}$. 则对任意的 t、$\tau\in[0,\ \infty)$，x_τ的 LB 估计 $\breve{x}_{\tau|t}$ 恒存在，且 $\breve{x}_{\tau|t}$ 及其均方误差阵 $Q_{\tau|t}$ 分别由下式算出：

$$\tilde{x}_{\tau|t}=\hat{x}_{\tau|t}+L_{\tau t}[\Gamma_{ap}S(K)_t(S(K)_t+S(K)_t\Gamma_{ap}S(K)_t)^-(Z(K)_t-S(K)_t\mu_{ap})+\mu_{ap}], \tag{7.3.12}$$

$$Q_{\tau|t}=P_{\tau|t}+L_{\tau t}[\Gamma_{ap}-\Gamma_{ap}S(K)_t(S(K)_t+S(K)_t\Gamma_{ap}S(K)_t)^-S(K)_t\Gamma_{ap}]L^*_{\tau t}. \tag{7.3.13}$$

特别是若 Γ_{ap} 为可逆阵，则以上两式可简化成

$$\tilde{x}_{\tau|t}=\hat{x}_{\tau|t}+L_{\tau t}(\Gamma_{ap}^{-1}+S(K)_t)^{-1}(z(K)_t+\Gamma_{ap}^{-1}\mu_{ap}), \tag{7.3.12'}$$

$$Q_{\tau|t}=P_{\tau|t}+L_{\tau t}(\Gamma_{ap}^{-1}+S(K)_t)^{-1}L^*_{\tau t}. \tag{7.3.13'}$$

其中 $\hat{x}_{\tau|t}$、$P_{\tau|t}$ 由定理 7.2.5 算出(但那里的 Ev_t、Ew_t 都用 0 代替)，$z(K)_t$、$S(K)_t$ 由定理 7.3.2 算出，$L_{\tau t}$ 由定理 7.3.3 算出.

证　由定理 2.3.5 与定理 2.2.6 立即证得. ▌

最后，考虑如下与(7.2.1)稍有不同的线性随机系统:

$$\begin{cases} x_t=\int_0^t A_s x_s ds+v_t, \\ y_t=\int_0^t J_s y_s ds+w_t \\ z_t=H_t x_t+y_t, \ (t\in[0,\ \infty)), \end{cases} \tag{7.3.14}$$

其中仍设 $u_{[0,\infty)}\triangleq\begin{pmatrix} v_{[0,\infty)} \\ w_{[0,\infty)} \end{pmatrix}$ 为 $n+m$ 维正交增量过程，且其核有(7.2.2)的形式. 又设对任意的 $t\in[0,\ \infty)$，有下式成立:

$$\int_0^t\|A_s\|^2ds<\infty,\quad \int_0^t\|J_s\|^2ds<\infty.$$

对于 H，假设它是绝对连续的 $C^{m\times n}$ 值函数，且其导函数 $\dot{H}$ 满足

$$\int_0^t\|\dot{H}_s\|^2ds<\infty.$$

对比(7.2.1)与(7.3.14)两个系统，其主要的不同点是：前者的量测 z_t 是状态 x_t 的某种"积累"(即积分)，而后者则是对 x_t 的"实时"量测. 另外，$y_{[0,\infty)}$ 也可看作是量测噪声，但它不是正交增量过程，而是由一个线性随机微分方程确定的. 这种形式的系统在工程技术文献中也是较常见的. 我们姑且称之为非积累量测系统，它与第 4 章§4.3 中讨论的带 ARMA 量测噪声的系统(离散

时间)有某种类似之处.

下面两个定理表明系统 (7.3.14) 可以等价于一个有(7.2.1)形式的系统.

定理 7.3.5 设 $z_{[0,\infty)}$ 为一二阶连续的 m 维局部正则过程:

$$\bar{z}_t \triangleq z_t - \int_0^t J_s z_s ds, \quad (t \in [0, \infty)).$$

则对任意的 $t \in [0, \infty)$ 有 $\mathscr{H}(\bar{z}, t) = \mathscr{H}(\bar{z}, t)$. 又对任一在$[0, t]$上连续的 $C^{k\times m}$ 值函数 F, 记

$$\bar{F}_s \triangleq F_s - \int_0^s F_r J_r^* dr, \quad (s \in [0, t]).$$

则 $F \in \mathscr{R}(z, t)^k$ 的必要与充分条件为 $\bar{F} \in \mathscr{R}(z, t)^k$. 且有 $z(F)_t = \bar{z}(\bar{F})_t$ 成立.

证　显然有 $\mathscr{H}(\bar{z}, t) \subset \mathscr{H}(z, t)$. 反之, 因 $\bar{z}_{[0,\infty)}$ 满足如下的线性随机微分方程:

$$z_t = \int_0^t J_s z_s ds + \bar{z}_t, \quad (t \in [0, \infty)).$$

由定理 7.1.2, 方程的解是唯一的. 由此知 $\mathscr{H}(z, t) \subset \mathscr{H}(\bar{z}, t)$.

现在设 $F \in \mathscr{R}(z, t)^k$, 由定理 6.1.3, F 必在$[0, t]$上连续. 又由已证的事实, 必有 $z(F)_t \in \mathscr{H}(\bar{z}, t)^k$, 且对任意的 $s \in [0, t]$ 有下式成立:

$$R(z(F)_t, \bar{z}_s) = R\left(z(F)_t, z_s - \int_0^s J_r z_r dr\right)$$

$$= F_s - \int_0^s F_r J_r^* dr = \bar{F}_s.$$

因此 $\bar{F} \in \mathscr{R}(\bar{z}, t)^k$, 且 $\bar{z}(\bar{F})_t = z(F)_t$. 反之, 设 $\bar{F} \in \mathscr{R}(z, t)^k$. 令 F 是如下方程的唯一解(见引理 7.1.3):

$$F_s = \int_0^s F_r J_r^* dr + \bar{F}_s, \quad (s \in [0, t]).$$

则 $\bar{z}(\bar{F})_t \in \mathscr{H}(z, t)^k$, 且有

$$R(\bar{z}(\bar{F})_t, z_s) = R\left(\bar{z}(\bar{F})_t, \int_0^s J_r z_r dr + \bar{z}_s\right)$$

$$-\int_0^s R(\bar{z}(\bar{F})_t,\ z_r)J_r^* dr+F_s,\quad (s\in[0,\ t]).$$

由解的唯一性即知有 $R(\bar{z}(\bar{F})_t,\ z_s)=F_s$. 因此 $F\in\mathscr{R}(z,\ t)^k$, 且 $z(F)_t=\bar{z}(\bar{F})_t$. ▌

定理 7.3.6 考虑系统(7.3.14), 设其满足所假定的条件. 对任意的 $t\in[0,\ \infty)$, 令

$$\bar{z}_t\triangleq z_t-\int_0^t J_s z_s ds,$$

$$\bar{w}_t\triangleq w_t+H_0 v_0+\int_0^t H_s dv_s.$$

又令

$$\bar{H}\triangleq\dot{H}+HA-JH,$$

$$\bar{D}\triangleq D+HCH^*+HB+B^*H,$$

$$\bar{B}\triangleq B+CH^*,$$

其中 C、D、B 见(7.2.2)式. 则可表为

$$\bar{z}_t=\int_0^t \bar{H}_s x_s ds+\bar{w}_t,$$

其中 $\bar{u}_{[0,\infty)}\triangleq\begin{pmatrix} v_{[0,\infty)} \\ \bar{w}_{[0,\infty)}\end{pmatrix}$ 为 $n+m$ 维正交增量过程, 且其核为

$$R(\bar{u}_t)=R(\bar{u}_0)+\int_0^t\begin{pmatrix} C_s & \bar{B}_s \\ \bar{B}_s^* & \bar{D}_s\end{pmatrix}ds.$$

证 由定理 7.1.5、定理 5.2.3 和分部积分公式容易算出:

$$\int_0^t(\dot{H}_s+H_sA_s)x_s ds$$

$$=\int_0^t(\dot{H}_s+H_sA_s)\left[\Phi(s,\ 0)x_0+\int_0^s\Phi(s,\ r)dv_r\right]ds$$

$$=\int_0^t\left(\dot{H}_s\Phi(s,\ 0)+H_s\frac{d\Phi(s,\ 0)}{ds}\right)ds\,x_0$$

$$+\int_0^t\left[\int_r^t\left(\dot{H}_s\Phi(s,\ r)+H_s\frac{\partial\Phi(s,\ r)}{\partial s}\right)ds\right]dv_r$$

$$=(H_t\Phi(t,\ 0)-H_0)x_0+\int_0^t(H_t\Phi(t,\ r)-H_r)dv_r$$

$$=H_t x_t-H_0 v_0-\int_0^t H_r dv_r.$$

由此及(7.3.14)即有

$$
\begin{aligned}
\bar{z}_t \triangleq z_t - \int_0^t J_s z_s ds
&= H_t x_t + y_t - \int_0^t J_s (H_s x_s + y_s) ds \\
&= \int_0^t (\dot{H}_s + H_s A_s) x_s ds + H_0 v_0 + \int_0^t H_r dv_r \\
&\quad + y_t - \int_0^t J_s (H_s x_s + y_s) ds \\
&= \int_0^t (\dot{H}_s + H_s A_s - J_s H_s) x_s ds + H_0 v_0 \\
&\quad + \int_0^t H_r dv_r + w_t \\
&= \int_0^t \overline{H}_s x_s ds + \overline{w}_t.
\end{aligned}
$$

其余部分由 $\overline{w}_{[0,\infty)}$ 的定义及(7.2.2)式易证得. ▌

根据上面两个定理，我们把系统(7.3.14)化成了如下与之等价的系统，而它有和系统(7.2.1)完全同样的形式：

$$
\begin{cases}
x_t = \int_0^t A_s x_s ds + v_t, \\
\bar{z}_t = \int_0^t \overline{H}_s x_s ds + \overline{w}_t.
\end{cases}
$$

于是，从定理7.2.2，我们就可以定出系统(7.3.14)输出的再生表示. 同样，从定理7.2.5、7.3.3和7.3.4，可以分别解决它的各种滤波问题.

第 8 章

线性随机系统的最优控制

本章的主要结果(定理 8.1.4 和 定理 8.2.4)是经典的，但文献中已有的证明要用到比较复杂或高深的工具．我们将在本章给出它们的相当简单而严格的证明；同时，也是为了把它们与滤波的结果作一对比．

§ 8.1 离散时间受控系统

定义 8.1.1 我们限于考虑白噪声系统．一个离散时间受控线性系统由如下的递推方程确定：

$$\begin{cases} x_t = \Phi_{t,t-1}x_{t-1} + \Gamma_t y_{t-1} + v_t; \\ z_t = H_t x_t + w_t, \end{cases} \quad (t \in N) \tag{8.1.1}$$

(依约定，当某值的下标 $t<0$ 时，总认定该值为 0)．其中 $\begin{pmatrix} v_N \\ w_N \end{pmatrix}$为 $n+m$ 维局部正则的正交随机序列(即白噪声)，并记

$$R\left(\begin{pmatrix} v_t \\ w_t \end{pmatrix}\right) \triangleq \begin{pmatrix} V_t & U_t \\ U_t^* & W_t \end{pmatrix}, \quad t \in N.$$

它与通常的线性系统(4.1.1)的不同之处，除了因为要与连续时间系统对照而稍微改动了一下记号外，主要是在动态方程中加入了"控制项"$\Gamma_t y_{t-1}$．其中 y_N 为一 k 维局部正则序列，称为控制策略，y_t 称为 t 时的控制变量，$\Gamma_t(\in C^{n\times k})$称为 t 时的控制系数矩阵．

显然，一个合理的控制策略应该是：它在 t 时给出的控制变量

只基于到 t 为止的系统的量测 z_{N_t}，而不依赖于系统的未来量测；而所谓线性控制，则是指 t 时的控制变量应是基于 z_{N_t} 构成的线性统计量，这句话用关系式来表达，就是对任意的 $t\in N$，有

$$y_t\in \mathscr{H}^+(z,\ t)^k\triangleq \mathscr{H}(z,\ t)^k+C^k. \tag{8.1.2}$$

满足这一条件的控制策略就称为容许线性控制策略。今后我们将只考虑这样的控制策略，并把条件 (8.1.2) 简记为 $y_N\in \mathscr{H}_N^{+k}$。如果仅限于讨论到某一时刻 τ 为止的控制，则简记为 $y_{N_{\tau-1}}\in \mathscr{H}_{N_{\tau-1}}^{+k}$。

像(8.1.1)这样的控制系统模型，在工程技术以致经济管理等领域的控制问题中是常见的，或可近似地归结为这类模型.

定理 8.1.2 对于受控系统(8.1.1)，x_t 基于量测的一步预测 $\hat{x}_{t|t-1}$，滤波 $\hat{x}_t(\triangleq\hat{x}_{t|t})$及其误差方差阵 $P_{t|t-1}$，$P_t(\triangleq P_{t|t})$有如下的递推公式：

$$\hat{x}_{t|t-1}=\Phi_{t,t-1}\hat{x}_{t-1}+\Gamma_t y_{t-1}+Ev_t,$$
$$P_{t|t-1}=\Phi_{t,t-1}P_{t-1}\Phi_{t,t-1}^*+V_t,$$
$$\varepsilon_t=z_t-H_t\hat{x}_{t|t-1}-Ew_t,$$
$$R(\varepsilon_t)=H_tP_{t|t-1}H_t^*+H_tU_t+U_t^*H_t^*+W_t,$$
$$\hat{x}_t=\hat{x}_{t|t-1}+(P_{t|t-1}H_t^*+U_t)R(\varepsilon_t)^-\varepsilon_t,$$
$$P_t=P_{t|t-1}-(P_{t|t-1}H_t^*+U_t)R(\varepsilon_t)^-(H_tP_{t|t-1}+U_t^*).$$

证 这里列举的公式与系 4.1.3 前一部分的公式几乎完全相同，除了对记号作了一点改动之外，就是在第一个公式中添加了 $\Gamma_t y_{t-1}$。根据条件(8.1.2)，这一项属于 $\mathscr{H}^+(z,\ t-1)^k$。因此由定理 2.1.3 有

$$\begin{aligned}\hat{x}_{t|t-1}&\triangleq \pi(x_t|z_{N_{t-1}})\\&=\pi(\Phi_{t,t-1}x_{t-1}+\Gamma_t y_{t-1}+v_t|z_{N_{t-1}})\\&=\Phi_{t,t-1}\hat{x}_{t-1}+\Gamma_t y_{t-1}+Ev_t,\end{aligned}$$

其余都是明显的. ∎

这里要特别指出的是：从公式可以看出，误差方差阵 $P_{t|t-1}$、P_t 以及 $R(\varepsilon_t)$ 都与容许控制策略 y_N 的选取无关.

下面着重讨论最优控制策略的选取问题. 为此，首先必须确

定一个策略好坏的标准.

定义 8.1.3 给定一个$(n+h)$阶半正定 Hermite 阵列:

$$S_t \triangleq \begin{pmatrix} M_t & L_t \\ L_t^* & N_t \end{pmatrix} \geqslant 0, \quad (t \in \boldsymbol{N}). \tag{8.1.3}$$

任意固定$\tau \in \boldsymbol{N}$, 定义

$$\beta(y, \tau) \triangleq \sum_{t=1}^{\tau} E\left[(x_t^* - Ex_t^*, y_{t-1}^* - Ey_{t-1}^*)\begin{pmatrix} M_t & L_t \\ L_t^* & N_t \end{pmatrix} \cdot \begin{pmatrix} x_t - Ex_t \\ y_{t-1} - Ey_{t-1} \end{pmatrix}\right] \tag{8.1.4}$$

则β称为系统(8.1.1)在$\boldsymbol{N}_\tau$上的一个平均平方损失函数. 其中S_t称为t时的损失加权矩阵, 它们可以由设计者根据需要加以确定和调整. 平方损失的实际意义是: 当状态x_t和控制变量y_{t-1}偏离它们的均值时, 必然会造成某种损失或增加费用, 而其诸分量损失的大小及它们之间的相互关系, 则可依据实际情况由加权矩阵诸元的大小反映出来. 当然, 还可能定义其他形式的损失函数. 但在线性控制理论中, 用得最多的还是如(8.1.4)式这样的平均平方损失函数. 除了它有明显的实际意义之外, 还因为从它可以完全解决最优控制的确定问题.

一个局限于$\boldsymbol{N}_\tau$上的容许线性控制策略$y_{N_{\tau-1}}^0$称为最优线性控制策略, 如果它使平均平方损失函数达到最小值, 即满足$y_{N_{\tau-1}}^0 \in \mathscr{H}_{N_{\tau-1}}^{+h}$, 且

$$\beta(y^0, \tau) = \inf\{\beta(y, \tau) : y_{N_{\tau-1}} \in \mathscr{H}_{N_{\tau-1}}^{+h}\}. \tag{8.1.5}$$

下面的定理给出了$y_{N_{\tau-1}}^0$的存在性及其构造方法.

定理 8.1.4 设系统(8.1.1)的平均平方损失函数由(8.1.3)和(8.1.4)式给出, 且加权子矩阵N_t满足条件$\Gamma_t N_t^- N_t = \Gamma_t$, $t \in \boldsymbol{N}_\tau$ (如N_t恒为可逆阵, 则自然满足). 又设n阶阵列Q_t $(t=\tau, \tau-1, \cdots, 2)$由如下的逆转递推方程算出:

$$Q_\tau = M_\tau, \quad \Delta_t = \Gamma_t^* Q_t \Gamma_t + \Gamma_t^* L_t + L_t^* \Gamma_t + N_t, \tag{8.1.6}$$

$$Q_{t-1|t} = Q_t - (Q_t \Gamma_t + L_t) \Delta_t^- (\Gamma_t^* Q_t + L_t^*), \tag{8.1.7}$$

$$Q_{t-1} = \Phi_{t,t-1}^* Q_{t-1|t} \Phi_{t,t-1} + M_{t-1}, \quad (t=\tau, \tau-1, \cdots, 2). \tag{8.1.8}$$

则 N_τ 上的最优线性控制策略为

$$y_{t-1}^0 = -\Delta_t^-(\Gamma_t^* Q_t + L_t^*)\Phi_{t,t-1}\hat{x}_{t-1}, \quad (t=2, \cdots, \tau). \quad (8.1.9)$$

其中滤波值 $\hat{x}_{t-1}$ 由定理 8.1.2 递推计算.

证 在证明过程中, 不失一般性, 不妨假定系统 (8.1.1) 中 v_N、w_N 和 y_N 的均值函数都等于 0, 因而 x_N、z_N 的均值函数也为 0, 否则, 可考虑减去均值后的等价系统.

我们首先确定最优的终端控制变量 $y_{\tau-1}^0$, 因为它仅对终端状态 x_τ 发生影响, 从而只影响 $\beta(y, \tau)$ 最后一项的值, 即

$$\begin{aligned}\alpha_\tau &\triangleq E\left[(x_\tau^*, y_{\tau-1}^*)\begin{pmatrix}M_\tau & L_\tau \\ L_\tau^* & N_\tau\end{pmatrix}\begin{pmatrix}x_\tau \\ y_{\tau-1}\end{pmatrix}\right] \\ &= E(x_\tau^* M_\tau x_\tau) + E(x_\tau^* L_\tau y_{\tau-1}) + E(y_{\tau-1}^* L_\tau^* x_\tau) \\ &\quad + E(y_{\tau-1}^* N_\tau y_{\tau-1}).\end{aligned}$$

由(8.1.1)及其假设条件容易算出:

$$\begin{aligned}&E(x_\tau^* M_\tau x_\tau) \\ &\quad = E[(\Phi_{\tau,\tau-1}x_{\tau-1} + \Gamma_\tau y_{\tau-1} + v_\tau)^* M_\tau(\Phi_{\tau,\tau-1}x_{\tau-1} \\ &\qquad + \Gamma_\tau y_{\tau-1} + v_\tau)] \\ &\quad = E(x_{\tau-1}^*\Phi_{\tau,\tau-1}^* M_\tau \Phi_{\tau,\tau-1}x_{\tau-1}) \\ &\qquad + E(x_{\tau-1}^*\Phi_{\tau,\tau-1}^* M_\tau \Gamma_\tau y_{\tau-1}) \\ &\qquad + E(y_{\tau-1}^*\Gamma_\tau^* M_\tau \Phi_{\tau,\tau-1}x_{\tau-1}) \\ &\qquad + E(y_{\tau-1}^*\Gamma_\tau^* M_\tau \Gamma_\tau y_{\tau-1}) + E(v_\tau^* M_\tau v_\tau), \\ &E(x_\tau^* L_\tau y_{\tau-1}) \\ &\quad = E[(\Phi_{\tau,\tau-1}x_{\tau-1} + \Gamma_\tau y_{\tau-1} + v_\tau)^* L_\tau y_{\tau-1}] \\ &\quad = E(x_{\tau-1}^*\Phi_{\tau,\tau-1}L_\tau y_{\tau-1}) + E(y_{\tau-1}^*\Gamma_\tau^* L_\tau y_{\tau-1}).\end{aligned}$$

将这两个结果代入 α_τ 式中, 并注意 $Q_\tau = M_\tau$ 及(8.1.6)式, 就有

$$\begin{aligned}\alpha_\tau &= E[y_{\tau-1}^*(\Gamma_\tau^* M_\tau \Gamma_\tau + \Gamma_\tau^* L_\tau + L_\tau^*\Gamma_\tau + N_\tau)y_{\tau-1}] \\ &\quad + E[y_{\tau-1}^*(\Gamma_\tau^* M_\tau + L_\tau^*)\Phi_{\tau,\tau-1}x_{\tau-1}] \\ &\quad + E[x_{\tau-1}^*\Phi_{\tau,\tau-1}^*(M_\tau\Gamma_\tau + L_\tau)y_{\tau-1}] \\ &\quad + E(x_{\tau-1}^*\Phi_{\tau,\tau-1}^* M_\tau \Phi_{\tau,\tau-1}x_{\tau-1}) + E(v_\tau^* M_\tau v_\tau) \\ &= E(y_{\tau-1}^*\Delta_\tau y_{\tau-1}) + E[y_{\tau-1}^*(\Gamma_\tau^* Q_\tau + L_\tau^*)\Phi_{\tau,\tau-1}x_{\tau-1}]\end{aligned}$$

$$+E[x_{\tau-1}^*\Phi_{\tau,\tau-1}^*(Q_\tau\Gamma_\tau+L_\tau)y_{\tau-1}]$$
$$+E(x_{\tau-1}^*\Phi_{\tau,\tau-1}^*Q_\tau\Phi_{\tau,\tau-1}x_{\tau-1})+E(v_\tau^*Q_\tau v_\tau).$$

在最后一式中,只有前三项含有 $y_{\tau-1}$,将它们配成"平方式",并注意(8.1.7)式,即得

$$\alpha_\tau=E\{[y_{\tau-1}+\Delta_\tau^-(\Gamma_\tau^*Q_\tau+L_\tau^*)\Phi_{\tau,\tau-1}x_{\tau-1}]^*\Delta_\tau[\cdots]\}$$
$$+E(x_{\tau-1}^*\Phi_{\tau,\tau-1}^*Q_{\tau-1,\tau}\Phi_{\tau,\tau-1}x_{\tau-1})+E(v_\tau^*Q_\tau v_\tau).\tag{8.1.10}$$

上式中[…]简单表示其中的内容与 Δ_τ 前的方括号相同. 此外还用到了关系式:

$$(Q_t\Gamma_t+L_t)\Delta_t^-\Delta_t=Q_t\Gamma_t+L_t.$$

这容易从假设条件 $\Gamma_tN_t^-N_t=\Gamma_t$ 及 $L_tN_t^-N_t=L_t$(由加权阵 S_t 的半正定性所蕴含)推出来.

为了定出最优终端控制 $y_{\tau-1}^0$, 还须注意两件事实: 一是如果记 $\tilde{x}_{\tau-1}\triangleq x_{\tau-1}-\hat{x}_{\tau-1}$,则由滤波的定义知 $\hat{x}_{\tau-1}\in\mathscr{H}^+(z,\tau-1)^n$, $\tilde{x}_{\tau-1}$ 与 $\hat{x}_{\tau-1}$ 及 $y_{\tau-1}$ 正交, 且 $E(\tilde{x}_{\tau-1}\tilde{x}_{\tau-1}^*)=P_{\tau-1}$. 二是由矩阵迹的性质 $\mathrm{tr}(AB)=\mathrm{tr}(BA)$知, 对任意两个二阶随机向量 x 与 y 有

$$E(x^*Ay)=E[\mathrm{tr}(x^*Ay)]=\mathrm{tr}[AE(yx^*)]$$

成立. 于是将 $x_{\tau-1}=\hat{x}_{\tau-1}+\tilde{x}_{\tau-1}$ 代入(8.1.10)式,并根据上述事实,则有

$$\alpha_\tau=\mathrm{tr}[\Delta_\tau R(y_{\tau-1}+\Delta_\tau^-(\Gamma_\tau^*Q_\tau+L_\tau^*)\Phi_{\tau,\tau-1}\hat{x}_{\tau-1})]$$
$$+\mathrm{tr}[\Phi_{\tau,\tau-1}^*(Q_\tau\Gamma_\tau+L_\tau)\Delta_\tau^-(\Gamma_\tau^*Q_\tau+L_\tau^*)\Phi_{\tau,\tau-1}P_{\tau-1}]$$
$$+\mathrm{tr}(Q_\tau V_\tau)+E(x_{\tau-1}^*\Phi_{\tau,\tau-1}^*Q_{\tau-1,\tau}\Phi_{\tau,\tau-1}x_{\tau-1}).\tag{8.1.11}$$

上式右边的四项都取非负实值, 且只有第一项与 $y_{\tau-1}$ 有关. 因此,为了使 α_τ 达到最小值, 只有取 $y_{\tau-1}$ 使第一项等于0. 也就是说,应有

$$y_{\tau-1}^0=-\Delta_\tau^-(\Gamma_\tau^*Q_\tau+L_\tau^*)\Phi_{\tau,\tau-1}\hat{x}_{\tau-1}\in\mathscr{H}^+(z,\tau-1)^k.$$

现在我们考虑倒数第二个控制变量 $y_{\tau-2}$. 显然, 它只影响 $B(y,\tau)$的最后两项之和. 但最后一项的值 α_τ 已经表成(8.1.11)

式，而且已取 $y_{\tau-1}=y^0_{\tau-1}$. 由此，再注意(8.1.8)式，这最后两项之和又可表成

$$\begin{aligned}(x^*_{\tau-1},\ y^*_{\tau-2})\begin{pmatrix} M_{\tau-1} & L_{\tau-1} \\ L^*_{\tau-1} & N_{\tau-1}\end{pmatrix}\begin{pmatrix} x_{\tau-1} \\ y_{\tau-2}\end{pmatrix}+\alpha_\tau \\ =(x^*_{\tau-1},\ y^*_{\tau-2})\begin{pmatrix} Q_{\tau-1} & L_{\tau-1} \\ L^*_{\tau-1} & N_{\tau-1}\end{pmatrix}\begin{pmatrix} x_{\tau-1} \\ y_{\tau-2}\end{pmatrix}+\lambda_\tau \\ \triangleq \alpha_{\tau-1}+\lambda_\tau,\end{aligned}$$

其中

$$\begin{aligned}\lambda_\tau \triangleq \operatorname{tr}[\Phi^*_{\tau,\tau-1}(Q_\tau\Gamma_\tau+L_\tau)\Delta^-_\tau(\Gamma^*_\tau Q_\tau+L^*_\tau)\Phi_{\tau,\tau-1}P_{\tau-1}] \\ +\operatorname{tr}(Q_\tau V_\tau)\end{aligned}$$

是与控制变量无关的常值(由定理 8.1.2 已知 P_t 与控制无关). 因此为求 $y^0_{\tau-2}$ 必须且只须控 $\alpha_{\tau-1}$ 达到极小值. 对比 $y^0_{\tau-1}$ 的求法，这时不过是把指标 τ 控成 $\tau-1$ 而已. 于是完全类似地可以求出

$$y^0_{\tau-2}=-\Delta^-_{\tau-1}(\Gamma^*_{\tau-1}Q_{\tau-1}+L^*_{\tau-1})\Phi_{\tau-1,\tau-2}\hat{x}_{\tau-2}.$$

照这样一步步递推下去(实际上是运用逆转归纳法)，即可证得定理的结论(8.1.9)式. 由证明过程顺便还可算出 N_τ 上的最小平方损失

$$\begin{aligned}&\beta(y^0,\ \tau) \\ &=\sum_{t=1}^{\tau}\{\operatorname{tr}[\Phi^*_{t,t-1}(Q_t\Gamma_t+L_t)\Delta^-_t(\Gamma^*_t Q_t+L^*_t)\Phi_{t,t-1}P_{t-1}] \\ &\quad+\operatorname{tr}(Q_t V_t)\}.\ \blacksquare\end{aligned}$$

注记 8.1.5 从重要的定理 8.1.4 中，显然可以看出如下两点：

i) 如果系统(8.1.1)的状态是可以完全量测的，即 $z_t=x_t$，($t\in[0,\ \infty)$)，那么(8.1.9)式中的 $\hat{x}_{t-1}$ 就可以换成 x_{t-1}. 或者反过来说，如果有了状态完全可量测时的最优线性控制公式，那么对于不完全量测的情形，只须把公式中的状态变量换成相应的滤控值即可. 这一事实在控制理论中有一个著名的名称，称为控制与滤波的“分离原理”(对非线性控制未必正确).

ii) 定理 8.1.4 与定理 8.1.2 中的诸递推公式存在明显的对

称性：$\Phi^*_{t,t-1}$ 与 $\Phi_{t,t-1}$，Γ^*_t 与 H_t，M_t 与 V_t，N_t 与 W_t，L_t 与 U_t，Q_t 与 $P_{t|t-1}$，$Q_{t-1|t}$ 与 P_t，Δ_t 与 $R(\varepsilon_t)$ 等形成一一对应，前者是指标由大到小(从后往前)递推的，而后者是指标由小到大(从前往后)递推的．这一现象在控制理论中也有一个著名的名称，称为控制与滤波的"对偶原理"．

§8.2 连续时间受控系统

定义 8.2.1 一个连续时间受控系统由如下的线性随机微分方程组确定：

$$\begin{cases} x_t = \int_0^t A_s x_s ds + \int_0^t G_s y_s ds + v_t, \\ z_t = \int_0^t H_s x_s ds + w_t, \end{cases} \quad (t \in [0, \infty)). \tag{8.2.1}$$

其中 $\begin{pmatrix} v_{[0,\infty)} \\ w_{[0,\infty)} \end{pmatrix}$ 为 $n+m$ 维(局部正则的)正交增量过程，且其核为

$$R\left(\begin{pmatrix} v_t \\ w_t \end{pmatrix}\right) = \begin{pmatrix} R(v_0) & R(v_0, w_0) \\ R(w_0, v_0) & R(w_0) \end{pmatrix} + \int_0^t \begin{pmatrix} C_s & B_s \\ B^*_s & D_s \end{pmatrix} ds$$
$$(t \in [0, \infty)).$$

A、G、H 分别是 $[0, \infty)$ 上 Borel 可测的 $C^{n\times n}$、$C^{n\times k}$、$C^{m\times n}$ 值函数，$DD^-H = H$，且对任意 $t \in [0, \infty)$ 满足

$$\int_0^t \|A_s\|^2 ds < \infty \quad \int_0^t \|G_s\|^2 ds < \infty,$$

$$\int_0^t \operatorname{tr}(H^*_s D^-_s H_s) ds < \infty.$$

它与通常的线性随机系统(7.2.1)的不同之处，在于动态方程中加入了控制项 $\int_0^t G_s y_s ds$．其中 $y_{[0,\infty)}$ 为一二阶可测的 k 维局部正则过程，称为控制策略．y_t 称为 t 时的控制变量．

同离散时间系统一样，一个控制策略称为容许线性控制策略，

如果对任意 $t\in[0,\infty)$，有

$$y_t\in\mathscr{H}^+(z,t)^k\triangleq\mathscr{H}(z,t)^k+C^k. \qquad (8.2.2)$$

或简记为 $y_{[0,\infty)}\in\mathscr{H}^{+k}_{[0,\infty)}$. 今后我们将只考虑这样的策略. 如果仅限于讨论有限时间段 $[0,\tau]$ 上的控制，则简记为 $y_{[0,\tau]}\in\mathscr{H}^{+k}_{[0,\tau]}$.

定理 8.2.2 对于受控系统(8.2.1)，x_t 的滤波 $\hat{x}_t$ 及其误差方差阵 P_t 满足如下的微分方程：

$$\varepsilon_t=z_t-\int_0^t H_s\hat{x}_s ds-Ew_t,$$

$$\hat{x}_t=R(x_0,z_0)R(z_0)^-(z_0-Ez_0)+\int_0^t A_s\hat{x}_s ds+\int_0^t G_s y_s ds$$
$$+\int_0^t(P_sH_s^*+B_s)D_s^- d\varepsilon_s+Ev_t,\ (t\in[0,\infty)),$$
$$\dot{P}=AP+PA-(PH^*+B)D^-(HP+B^*)+C,$$
$$P_0=R(x_0)-R(x_0,z_0)R(z_0)^-R(z_0,x_0).$$

证 与定理 7.2.5 中的公式相比较，这里不同的只是在 $\hat{x}_t$ 的方程中添加了 $\int_0^t G_s y_s ds$ 一项，而根据(8.2.2)，它属于 $\mathscr{H}^+(z,t)^n$. 不失一般性，可假定 $v_{[0,\infty)}$ 与 $w_{[0,\infty)}$ 的均值函数为 0. 用 Φ 表示 A 的转移阵. 仿定理 7.1.5. 的 ii)的证明易知，(8.2.1)的前一方程的解可以表为

$$x_t=\Phi(t,0)x_0+\int_0^t\Phi(t,s)G_s y_s ds+\int_0^t\Phi(t,s)dv_s. \qquad (8.2.3)$$

于是有

$$\begin{aligned}\hat{x}_t&\triangleq\pi(x_t|z_{[0,t]})\\&=\pi\left(\Phi(t,0)x_0+\int_0^t\Phi(t,s)dv_s\Big|z_{[0,t]}\right)\\&\quad+\int_0^t\Phi(t,s)G_s y_s ds.\end{aligned}$$

但由定理 7.2.5 知道

$$\pi\left(\Phi(t,0)x_0+\int_0^t\Phi(t,s)dv_s\Big|z_{[0,t]}\right)$$

$$=\Phi(t,0)R(x_0,z_0)R(z_0)^{-}z_0$$
$$+\int_0^t \Phi(t,s)(P_sH_s^*+B_s)D_s^{-}d\varepsilon_s.$$

代入前一式，再表为随机微分方程的形式，即得所证. ▮

同离散时间系统一样，需要指出的是：滤波误差方差阵 P_t 与容许控制策略 $y_{[0,\infty)}$ 的选取无关.

为了研究连续时间受控系统的最优控制问题，下面定义该系统的损失函数。

定义 8.2.3 在 $[0,\infty)$ 上给定一个 Borel 可测且取值为 $(n+k)$ 阶半正定 Hermite 阵的函数

$$S_t \triangleq \begin{pmatrix} M_t & L_t \\ L_t^* & N_t \end{pmatrix} \geqslant 0,\ (t\in[0,\infty)). \tag{8.2.4}$$

任意固定 $\tau\in[0,\infty)$，设 $\|S\|$ 在 $[0,\tau]$ 上为有界. 又设 T_τ 为一 n 阶半正定 Hermite 阵. 定义

$$\beta(y,\tau)\triangleq\int_0^\tau E\left[(x_t^*-Ex_t^*,\ y_t^*-Ey_t^*)\begin{pmatrix} M_t & L_t \\ L_t^* & N_t \end{pmatrix}\cdot\begin{pmatrix} x_t-Ex_t \\ y_t-Ey_t \end{pmatrix}\right]dt$$
$$+E[(x_\tau^*-Ex_\tau^*)T_\tau(x_\tau-Ex_\tau)]. \tag{8.2.5}$$

则 β 称为系统(8.2.1)在 $[0,\tau]$ 上的一个平均平方损失函数. 其中 S_t *称为 t 时的损失加权矩阵，T_τ 称为终端损失加权矩阵.* 它们可以由设计者根据需要加以确定和调整，反映当状态和控制变量偏离它们的均值时所造成的损失或费用.

一个局限于 $[0,\tau]$ 上的容许线性控制策略 $y^0_{[0,\tau]}$ *称为最优线性控制策略*，如果它使平均平方损失函数达到最小值，即满足 $y^0_{[0,\tau]}\in\mathscr{H}^{+k}_{[0,\tau]}$ 且

$$\beta(y^0,\tau)=\inf\{\beta(y,\tau):y_{[0,\tau]}\in\mathscr{H}^{+k}_{[0,\tau]}\}.$$

下面是本节的主要定理.

定理 8.2.4 设系统(8.2.1)的平均平方损失函数由(8.2.4)和(8.2.5)给出，且加权子矩阵 N_t 满足 $G_tN_t^{-}N_t=G_t$ $(t\in[0,\tau])$

(若 N_t 恒为可逆阵，则自然满足)．又设 Q_t ($t\in[0,\tau]$)是如下 n 阶阵的 Ricatti 型微分方程的解：

$$\begin{cases} -\dot{Q}=A^*Q+QA-(QG+L)N^-(G^*Q+L^*)+M; \\ Q_\tau=T_\tau. \end{cases} \tag{8.2.6}$$

则 $[0,\tau]$ 上的最优线性控制策略为

$$y_t^0=-N_t^-(G_t^*Q_t+L_t^*)\hat{x}_t,\ (t\in[0,\tau]). \tag{8.2.7}$$

其中 $\hat{x}_t$ 是 x_t 的滤波值．

证　对于连续时间系统，我们不可能再运用定理 8.1.4 证明中的逆递推法，因而它的证明也不再那样具有"构造性"．

首先，利用矩阵迹的性质，可以把平均平方损失函数(8.2.5)改写成如下形式：

$$\begin{aligned}\beta(y,\tau)=&\int_0^\tau \operatorname{tr}(M_tR(x_t))dt+\operatorname{tr}(T_\tau R(x_\tau))\\&+\int_0^\tau \operatorname{tr}[N_tR(y_t)+R(y_t,x_t)L_t\\&+L_t^*R(x_t,y_t)]dt.\end{aligned} \tag{8.2.8}$$

由(8.2.3)式有

$$\begin{aligned}R(y_t,x_t)=&R(y_t,x_0)\Phi^*(t,0)+\int_0^t R(y_t,y_s)G_s^*\Phi^*(t,s)ds\\&+R\left(y_t,\int_0^t\Phi(t,s)dv_s\right),\\R(x_t)=&\Phi(t,0)R(x_0)\Phi^*(t,0)\\&+\int_0^t\Phi(t,s)G_sR(y_s,x_0)\Phi^*(t,0)ds\\&+\int_0^t\Phi(t,0)R(x_0,y_s)G_s^*\Phi^*(t,s)ds\\&+\int_0^t\int_0^t\Phi(t,s)G_sR(y_s,y_r)G_r^*\Phi^*(t,r)ds\,dr\\&+\int_0^t\Phi(t,s)G_sR\left(y_s,\int_0^t\Phi(t,r)dv_r\right)ds\\&+\int_0^t R\left(\int_0^t\Phi(t,r)dv_r,y_s\right)G_s^*\Phi^*(t,s)ds\end{aligned}$$

$$+\int_0^t \Phi(t,\ s)C_s\Phi^*(t,\ s)ds.$$

将后一式稍加整理并用前一式的结果代入，容易算出

$$R(x_t)=\Phi(t,\ 0)R(x_0)\Phi^*(t,\ 0)$$
$$+\int_0^t \Phi(t,\ s)(G_sR(y_s,\ x_s)+R(x_s,\ y_s)G_s^*$$
$$+C_s)\Phi^*(t,\ s)ds. \qquad (8.2.9)$$

其次，由转移阵 Φ 的性质：$\Phi(s,\ t)=\Phi^{-1}(t,\ s)$（定理 7.1.5 的 i)）和逆阵求导公式得：

$$\frac{d\Phi(s,\ t)}{dt}=\frac{d\Phi^{-1}(t,\ s)}{dt}=-\Phi^{-1}(t,\ s)\left(\frac{d\Phi(t,\ s)}{dt}\right)\Phi^{-1}(t,\ s)$$
$$=-\Phi(s,\ t)A_t\Phi(t,\ s)\Phi(s,\ t)=-\Phi(s,\ t)A_t.$$

由此容易直接验证方程 (8.2.6) 的唯一解 Q_t 必满足如下积分方程：

$$Q_t=\int_t^\tau \Phi^*(s,\ t)[M_s-(Q_sG_s+L_s)N_s^-(G_s^*Q_s+L_s^*)]\Phi(t,\ s)ds$$
$$+\Phi^*(\tau,\ t)T_\tau\Phi(\tau,\ t)$$
$$\triangleq\overline{Q}_t+\overline{T}_t,\ (t\in[0,\ \tau]). \qquad (8.2.10)$$

再令

$$\overline{M}_t\triangleq M_t-(Q_tG_t+L_t)N_t^-(G_t^*Q_t+L_t^*), \qquad (8.2.11)$$

利用矩阵迹的性质及交换积分次序，由(8.2.9)和(8.2.10)式可得

$$\int_0^\tau \mathrm{tr}(\overline{M}_tR(x_t))dt=\int_0^\tau \mathrm{tr}(\overline{M}_t\Phi(t,\ 0)R(x_0)\Phi^*(t,\ 0))dt$$
$$+\int_0^\tau \mathrm{tr}\Big[\overline{M}_t\Big(\int_0^t \Phi(t,\ s)(G_sR(y_s,\ x_s)+R(x_s,\ y_s)G_s^*$$
$$+C_s)\Phi^*(t,\ s)ds\Big]dt$$
$$=\mathrm{tr}\Big[\Big(\int_0^\tau \Phi^*(t,\ 0)\overline{M}_t\Phi(t,\ 0)dt\Big)R(x_0)\Big]$$
$$+\int_0^\tau \mathrm{tr}\Big[\Big(\int_s^\tau \Phi^*(t,\ s)\overline{M}_t\Phi(t,\ s)dt\Big)(G_sR(y_s,\ x_s)$$
$$+R(x_s,\ y_s)G_s^*+C_s)\Big]ds$$

$$=\operatorname{tr}(\bar{Q}_0R(x_0))+\int_0^\tau\operatorname{tr}(\bar{Q}_sC_s)ds$$
$$+\int_0^\tau\operatorname{tr}(R(y_s,\ x_s)\bar{Q}_sG_s+G_s^*\bar{Q}_s(R(x_s,\ y_s))ds.\qquad(8.2.12)$$

用类似的计算又可得

$$\operatorname{tr}(T_\tau R(x_\tau))=\operatorname{tr}(\bar{T}_0R(x_0))+\int_0^\tau\operatorname{tr}(\bar{T}_sC_s)ds$$
$$+\int_0^\tau\operatorname{tr}(R(y_s,x_s)\bar{T}_sG_s+G_s^*\bar{T}_sR(x_s,\ y_s))ds.$$
$$(8.2.13)$$

将(8.2.12)与(8.2.13)两式相加,再由(8.2.10)即有

$$\int_0^\tau\operatorname{tr}(\bar{M}_tR(x_t))dt+\operatorname{tr}(T_\tau R(x_\tau))$$
$$=\operatorname{tr}(Q_0R(x_0))+\int_0^\tau\operatorname{tr}(Q_sC_s)ds$$
$$+\int_0^\tau\operatorname{tr}(R(y_s,\ x_s)Q_sG_s+G_s^*Q_sR(x_s,\ y_s))ds.$$
$$(8.2.14)$$

将(8.2.14)代入(8.2.8)式,并注意(8.2.11),就得到如下重要的关系式:

$$\beta(y,\ \tau)=\operatorname{tr}(Q_0R(x_0))+\int_0^\tau\operatorname{tr}(Q_sC_s)ds$$
$$+\int_0^\tau\operatorname{tr}[(Q_sG_s+L_s)N_s^-(G_s^*Q_s+L_s^*)R(x_s)]ds$$
$$+\int_0^\tau\operatorname{tr}[R(y_s,\ x_s)(Q_sG_s+L_s)$$
$$+(G_s^*Q_s+L_s^*)R(x_s,\ y_s)]ds+\int_0^\tau\operatorname{tr}(N_sR(y_s))ds$$
$$=\operatorname{tr}(Q_0R(x_0))+\int_0^\tau\operatorname{tr}(Q_sC_s)ds$$
$$+\int_0^\tau\operatorname{tr}[N_sR(y_s+N_s^-(G_s^*Q_s+L_s^*)x_s)]ds.$$
$$(8.2.15)$$

最后一个等式用到了$(Q_sG_s+L_s)N_s^-N_s=Q_sG_s+L_s$.

剩下的部分与(8.1.11)式的推导颇为相似，令 $\tilde{x}_s \triangleq x_s - \hat{x}_s$，则由滤波的定义知 $\hat{x}_s \in \mathscr{H}^+(z, s)^n$，$\tilde{x}_s$ 与 $\hat{x}_s$ 及 y_s 正交，且 $R(\tilde{x}_s) = P_s$ 与控制策略的选取无关. 用 $x_s = \hat{x}_s + \tilde{x}_s$ 代入(8.2.15)式，由所述的性质即有

$$\beta(y, \tau) = \operatorname{tr}(Q_0 R(x_0)) + \int_0^\tau \operatorname{tr}(Q_s C_s) ds$$

$$+ \int_0^\tau \operatorname{tr}[(Q_s G_s + L_s) N_s^- (G_s^* Q_s + L_s^*) P_s] ds$$

$$+ \int_0^\tau \operatorname{tr}[N_s R(y_s + N_s^- (G_s^* Q_s + L_s^*) \hat{x}_s)] ds.$$

上式右边的四项都取非负值，且只有第四项与 $y_{[0,\tau]}$ 有关. 因此，为了使 $\beta(y, \tau)$ 达到最小值，只有取 $y_{[0,\tau]}$ 使第四项等于 0，即应取

$$y_s = -N_s^- (G_s^* Q_s + L_s^*) \hat{x}_s \quad (s \in [0, \tau]).$$

这正是(8.2.7)式. ▌

对连续时间的线性受控随机系统，显然也有如注记 8.1.5 所说的控制与滤波的“分离原理”与“对偶原理”. 在此不再赘述.

参 考 文 献

[1] J. Hájek: On linear statistical problem in stochastic processes, Чехословацкий мат. жур. т. 12(1962).

[2] T. Kailath: RKHS approach to detection and estimation problems, IEEE Trans., Part I, IT- 17(1971); Part II, IT- 21(1975).

[3] T. Kailath et al: An innovations approach to least -square estimation, IEEE Trans., Part I, Part II, AC -13(1968); Part III, AC-16(1971); Part IV, AC-18(1973).

[4] M. Loève: Probability theory, 4th edition, Springer-Verlag, 1977~1978.

[5] K. J. Aström: Introduction to stochastic control theory, Academic Press, NewYork, 1970.

[6] R. Sh. Liptzer, A. N. Shiryaev: Statistics of stochastic processes, Springer-Verlag, 1977(中译本:里普切尔、史里亚耶夫:《随机过程统计》,宇航出版社,1987).

[7] Y. A. Rozanov: Innovation and non-anticipative processes, Proceeding of the 3rd-International Symposium on Multivariate Analysis, Dayton Ohio, June 1972.

[8] 中国科学院数学研究所概率组:《离散时间系统滤波的数学方法》,国防工业出版社,1975.

[9] 严加安:《鞅与随机积分引论》,上海科学技术出版社,1981.

[10] 安鸿志、陈兆国、杜金观、潘一民:《时间序列的分析与应用》,科学出版社,1986.

[11] 杜金观、潘一民:"协方差可分离随机序列的新息与滤波",《数学学报》,20(1977).

[12] 杜金观、潘一民:"可分离随机过程的新息与滤波",《数学学报》,24 (1981).

[13] 潘一民:"关于正态测度的绝对连续性与奇异性",《数学进展》,(1966).

[14] 潘一民:"一类统计量的在线递推算法",《应用数学学报》,4 (1981).

[15] 陈翰馥:"缺初值估计的最优性",《中国科学》,(1978).

[16] 陈翰馥:《离散时间系统的递推估计与随机控制》,科学出版社,1980.

[17] 谢衷洁、程乾生:"在一类非平稳干扰下的极大信噪比线性滤波问题",《数学学报》,22 (1979).

[18] 谢衷洁、程乾生:"具有平稳干扰的极大信噪比滤波问题",《应用数学学报》,4 (1981).

[19] 汪嘉冈:"二阶过程的最大信噪比滤波",《数学学报》,28 (1985).

[20] 汪咬元:"状态方程与量测方程中均含有参向量时的递推滤波",《科学通报》(1985).

[21] A. P. Verbyla, W. N. Venables: An extension of growth curve model, Biometrika, 75, 1(1988).

名 词 索 引